



Stød og stemmekvalitet

En akustisk-fonetisk undersøgelse af ændringer i stemmekvaliteten i forbindelse med stød

Hansen, Gert Foget

Publication date:
2015

Document version
Tidlig version også kaldet pre-print

Document license:
[CC BY-NC-ND](#)

Citation for published version (APA):
Hansen, G. F. (2015). *Stød og stemmekvalitet: En akustisk-fonetisk undersøgelse af ændringer i stemmekvaliteten i forbindelse med stød*. Det Humanistiske Fakultet, Københavns Universitet.

Stød og stemmekvalitet

*En akustisk-fonetisk undersøgelse
af ændringer i stemmekvaliteten
i forbindelse med stød*

ph.d.-afhandling

*Gert Foget Hansen
2015*

Stød og stemmekvalitet

Ph.d.-afhandling

Gert Foget Hansen

2015

Afdeling for Dialektforskning

Nordisk Forskningsinstitut

Det Humanistiske Fakultet

Københavns Universitet

Indhold

1 Indledning.....	7
2 Baggrundsviden.....	8
2.1 Fakta om stød.....	8
2.2 Nogle akustiske grundbegreber.....	10
2.3 Modulation: variation over temaet periodiske signaler.....	14
2.3.1 Amplitudemodulation.....	14
2.3.2 Frekvensmodulation.....	15
2.3.3 Modulation generelt.....	15
2.4 Stemmekildens funktionsmåde og akustiske egenskaber.....	16
2.4.1 Mellem modal stemthed og knirk: diplofoni.....	18
2.4.2 Glottis som støjkilde.....	19
2.4.3 Indkredsning af begrebet modal stemthed.....	19
3 Baggrund: tidligere undersøgelser af stød.....	20
3.1 Fischer-Jørgensen.....	22
3.1.1 Akustiske resultater.....	25
3.1.2 Fysiologiske resultater.....	30
3.1.3 Opsamling.....	35
3.1.4 Diskussion: stød as a phonation type:.....	36
3.1.5 Diskussion: variability and invariance.....	39
3.2 Grønnum & Basbøll (2001-2007).....	40
3.2.1 Vokalvarighed.....	41
3.2.2 Konsonantvarighed.....	42
3.2.3 Timing af stødfasen (laryngalisering).....	42
3.2.4 Varighedsforhold.....	44
3.2.5 Bemærkninger til varighedsforhold.....	44
3.2.6 Grønnum og Basbølls ballistiske model.....	45
3.2.7 Bemærkninger til Grønnum & Basbølls model.....	46
3.3 Knirk – fra uønsket epifænomen til definerende træk.....	52
4 Ny viden om knirk og en hypotese om stød.....	54
4.1 Stød fonetisk set – status quo.....	54
4.2 Knirk eller ej?.....	54
4.3 Forklaringsbehov I: Hvad er stød fonetisk set?.....	56
4.4 Forklaringsbehov II: Hvordan kan "synlige" og "usynlige" stød forekomme lige tydelige?.....	57
4.5 Ny forklaringsmodel for stød.....	58
4.6 Forbehold og implicitte antagelser.....	59
4.7 Forklaringspotentiale – muligheder og begrænsninger.....	60
4.8 Eksplicitering af forskelle mellem de to forklaringsmodeller.....	61
4.9 Kort om hvordan hypotesen kan testes.....	61
5 Måder at kvantificere (aspekter ved) stemmekvalitet.....	62
5.1 Kilde-filter-modellen.....	63

5.2 Stemmekilden – funktionsmåder og modelering.....	64
5.3 Akustisk analyse: beskrivelse af kilde og filterfunktionen.....	67
5.3.1 Akustisk analyse: estimering af filterfunktionen.....	67
5.3.2 Akustisk analyse: estimering af kildefunktionen – invers filtrering.....	67
5.4 Holistisk akustisk analyse : muligheder og problemer.....	69
5.4.1 Et eksempel: Interaktion mellem H1:H2 og F1.....	69
5.5 Periodicitet: yderligere udfordringer og muligheder.....	70
5.5.1 Udfordring for bestemmelse af H1:H2.....	70
5.5.2 Alternativer til H1:H2.....	70
5.5.3 Muligheder for beskrivelse af graden af periodicitet.....	71
5.5.4 Aperiodicitet udfordrer modeller for kildens funktionsmåde.....	72
5.5.5 Formulering af akustisk-fonetiske delhypoteser.....	73
6 Akustisk analyse – implementering og test af metoder.....	75
6.1 Oversigt over valgte metoder.....	76
6.2 Implementering og introduktion til den visuelle fremstilling.....	77
6.3 H1:H2.....	80
6.4 CoG (Centre of Gravity).....	81
6.5 BED (Band Energy Difference).....	81
6.6 Motivation for H1:H2, CoG og BED.....	82
6.7 H1:H1½ – et mål for diplofoni.....	84
6.8 Harmonicity to Noise Ratio (HNR).....	84
6.9 Motivation for HNR og H1:H1½.....	85
6.10 f0-tracking.....	86
6.10.1 To f0 estimerer.....	86
6.10.2 Visuel verifikation af f0.....	87
6.10.3 Udglatning og interpolation.....	88
6.11 Formanter.....	88
6.12 Intensitet.....	89
6.13 Afprøvning af de akustiske analyser på testmateriale.....	90
6.14 Test af filterbreddens indflydelse ved beregningen H1:H2.....	90
6.15 Afprøvninger af H1:H2, BED og H1:H1½ robusthed ifht. f0-fejl.....	93
6.15.1 H1:H2.....	93
6.15.2 BED.....	96
6.16 Test af overensstemmelse og afvigelser mellem H1:H2, BED og CoG.....	98
6.16.1 Serie 1, varierende stemmekvalitet, [i] og [a], middel grundtone.....	98
6.16.2 Serie 2, varierende grundtone, [i] og [a], middel kompression.....	100
6.16.3 Serie 3, varierende vokalkvaliteter, middel grundtone, middel kompression.....	103
6.17 Opsummering af test.....	106
6.18 Er metoderne hurtige nok?.....	108
6.19 Test af H1:H1½ – Afsøgning af diplofoni.....	110
6.20 Diplofoni: amplitude eller frekvensmodulation.....	114
7 Udformning af materiale.....	121
7.1 Generelle overvejelser.....	121

7.2	Oversigt over faktorer.....	122
7.3	Gennemgang af faktorer.....	127
7.3.1	Vokal vs. konsonant.....	127
7.3.2	Vokalkvalitet.....	127
7.3.3	Segmentelle omgivelser.....	127
7.3.4	Ordgrænse og åben vs. lukket stavelse.....	127
7.3.5	en- og to-stavelsesord.....	128
7.3.6	Hovedtryk vs. bitryk.....	128
7.3.7	Morfologisk struktur.....	130
7.3.8	Interaktioner.....	131
7.3.9	Ekstramateriale.....	131
7.4	Design – argumenter for oplæsning.....	132
7.5	Beskrivelse af sætningsmaterialet.....	132
7.6	Bemærkning om det undersøgte materiales omfang.....	132
7.7	Om optagelsen.....	133
8	Akustisk analyse af materialet.....	134
8.1	Gennemgang af eksempler på ordene ma'ler og maler.....	134
8.2	Præliminær sammenfatning på baggrund af de første eksempler.....	142
8.3	ma'ser vs. mases.....	143
8.4	ta'ber vs. taber.....	147
8.5	li'men vs. limet.....	152
8.6	lamm'et vs. lammet.....	156
8.7	Rekapitulering af hypotesen og de akustiske delhypoteser.....	162
8.7.1	Fire spørgsmål: specificering og eksempler.....	164
8.8	Delhypotese I.....	165
8.9	Delhypotese II.....	166
8.9.1	Positive eksempler.....	166
8.9.2	Negative eksempler.....	166
8.10	Delhypotese III og IV.....	167
8.10.1	Eksempler på tokens med en positiv evaluering.....	167
8.10.2	Eksempler der bedømmes negativt i forhold til spørgsmålet.....	168
8.11	Operationalisering af besvarelse af delspørgsmål.....	172
8.12	Svarkategorier.....	172
8.13	Selektering.....	173
8.14	Resultater – opdeling af materialet i fire undergrupper.....	175
8.15	Resultater for gruppe A.....	176
8.16	Resultater for gruppe B.....	177
8.17	Resultater for gruppe C.....	178
8.18	Resultater for gruppe D.....	180
9	Akustisk analyse: resultater.....	181
9.1	Evaluerings af hovedhypotese.....	181
9.2	Evaluerings af delhypotese II.....	182
9.3	Evaluerings af delhypotese IV.....	183
9.3.1	Opsamling på resultaterne for delhypotese II og IV.....	183

9.4 Observation af variation – to yderpunkter.....	184
9.5 Observation: tegn på at variationen ikke er tilfældigt fordelt.....	186
9.5.1 Opsamling på observationerne vedrørende variation i realisationen.....	187
9.6 Observation: diplofoni ser ud til at være udbredt.....	188
10 Afrunding.....	189
11 Referenceliste.....	190
Dansk resume.....	195
Summary in English.....	196

1 Indledning

Stød er en prosodi der knytter sig til visse stavelser. Stød er ofte betydningsadskillende, men distributionen følger i høj grad forudsigelige omend temmelig komplekse mønstre, se fx Basbøll 1985 og 1988. Fonetisk kan stød beskrives som "en uregelmæssighed i stemmelæbesvingningerne, en art knirkestemme" (Grønnum 2001, p. 45); "Stød is a kind of creaky voice" (Grønnum 1998, p. 103).

Det er imidlertid langt fra sådan at der altid ses uregelmæssige stemmelæbesvingninger i forbindelse med stød. I det hele taget har stød en meget varierende fremtoning. Selv i distinkt tale ser det ikke ud til at stød udviser konsekvente akustiske træk. Det er forunderligt, og det er det forhold der er baggrunden for denne undersøgelse.

Teksten er struktureret på denne måde: Først opsummeres resultaterne fra de væsentligste nyere fonetiske undersøgelser af stød, særligt Fisher-Jørgensen 1989a, og Grønnum & Basbøll 2001a,b, 2002a,b, 2003a,b og 2007. Variabiliteten i fremtoningen er et tilbagevendende tema i disse undersøgelser.

Der fremsættes en hypotese som ville kunne forklare en væsentlig del af variationen. Forskellige akustiske metoder der kan anvendes til at afprøve hypotesen diskuteres og testes. Hypotesen afprøves på et mindre materiale. Resultatet af afprøvningen falder overvejende negativt ud og hypotesen må derfor afvises. Arbejdet med at afprøve hypotesen giver dog ny viden om en række forhold.

Det viser sig at den sammenhæng mellem høj kompression og forekomst af uregelmæssige stemmelæbesvingninger man traditionelt har antaget ikke nødvendigvis gør sig gældende. De akustiske analyser viser at stød kan realiseres på to forskellige måder som begge indebærer uregelmæssige stemmelæbesvingninger. Forekomsten af de to typer er ikke ligeligt fordelt over det undersøgte materiale. Diplofoni (som er en form for regelmæssig variation i stemmelæbernes svingningsmønstre), ser ud til at forekomme lige så hyppigt som egentlig uregelmæssige stemmelæbesvingninger. Det ville derfor være ønskværdigt hvis diplofoni var mere velbeskrevet end tilfældet er.

2 Baggrundsviden

Før vi går til en mere detaljeret gennemgang af fonetiske undersøgelser af stød, vil det være formålstjenligt at præsentere nogle basale fakta om stød (2.1) og at give en kort introduktion til nogle vigtige akustiske grundbegreber (2.2 og 2.3) og til stemmekildens virkningsmåde (2.4).

2.1 Fakta om stød

Nedenstående oversigt opregner i punktform nogle væsentlige forhold der gør sig gældende for stød, og som danner rammen for denne undersøgelse. Oversigten tager afsæt i standardbeskrivelser af stød som det forekommer i ”Dansk rigsmål”¹.

1. Stød er en binær størrelse (dvs. vi opfatter ikke grader af stød).
2. Støds distribution er begrænset:
 - a. Der forekommer højst et stød pr. stavelse
 - b. Stød forekommer (med nogle specielle undtagelser) kun i stavelser med hoved eller bitryk. Hvis en stavelse mister sit tryk, tabes stødet også.
 - c. Stød kan kun forekomme i stavelser med såkaldt stødbasis – dvs. i stavelser med lang vokal eller med kort vokal fulgt af sonorant konsonant.
3. Manifestationen af stød er temmelig variabel. I akustiske analyser viser det sig som en kompleks variation af en række faktorer som vedrører bl.a. periodelængde (grundtone) og periodestabilitet (regelmæssig overfor uregelmæssig), frekvensspektret og intensitet.

Selvom stød set fra en akustisk og instrumentel synsvinkel kan udvise en kompleks variation både fra taler til taler og for hver taler for sig, så perciperes de ofte lige godt som stød. Auditivt kommer forskellene tilsyneladende kun til udtryk som mere eller mindre tydelige stød, typisk betegnet som kraftige eller svage.

Sammenhængen mellem de akustiske mål og den opfattede grad af distinkthed er ikke trivial: Stød som er vanskelige at registrere i akustiske analyser, kan sagtens være tydelige at høre, og stød som lyder kraftige, giver ikke nødvendigvis tydeligere udslag i de akustiske analyser.

Den variation som kan iagttages, både auditivt og akustisk, ser ikke ud til at følge en fast systematik, og må derfor betegnes som fri variation.

1 Både nomenklaturen og afgrænsningen af det dansk der inkluderes henholdsvis ekskluderes, varierer fra forfatter til forfatter, men der er et fælles fokus på dansk som det tales i den øvre middelklasse i København og omegn. Eksempler: Dansk Rigssprog (Smith 1944), Standard Danish (Fischer-Jørgensen 1989a), Moderne københavnsk rigsmål, eller Københavnsk rigsmål (fx Grønnum 2001)

Punkt 1, 2a og 2b er fonologisk funderede: Fonologiske analyser har ikke givet anledning til at operere med grader af stød. 2a og 2b er baseret på observeret distribution.

Punkt 2c er traditionelt set fonetisk funderet: Stødbasis beskrives som en fonetisk betingelse, og begrænsningen kommer af at stødet siges at kræve en vis mængde stemt materiale for at kunne udfolde sig (se fx Fisher-Jørgensen 1989a, p. 8)².

Punkt 3 vedrører den fonetiske realisering. Derudover er det måske værd at bemærke at selvom stød – som alle sproglyde – kan udtales mere eller mindre distinkt³, så er stød tilsyneladende bemærkelsesværdigt robust over for de reduktionsprocesser der optræder i talesproget, og så længe de berørte stavelers tryk ikke reduceres undergår stød tilsyneladende ikke reduktion eller bortfald selv under omstændigheder som ellers i øvrigt befordrer reduktion og bortfald af segmenter eller stavelser (Grønnum et. al 2013 og Schachtenhaufen 2013).

Hvor der på det fonologiske niveau hersker en ganske klar opfattelse af hvornår der kan tales om stød, så er der ikke på tilsvarende måde vished eller enighed om hvad stød er rent fonetisk set – hverken akustisk eller artikulatorisk.

Den eller de præcise artikulatoriske mekanismer bag realisationen af stød er ikke klar. Der er udbredt enighed om at stød er et glottalt fænomen, og resultaterne fra Fischer-Jørgensen 1989a peger ret tydeligt på at stød indebærer en glottal konstriktion. Der er dog samtidig enighed om at stød ikke generelt kan sidestilles med glottalt lukke, men at kraftige stød, fx under emfase kan realiseres som – eller indebære – et glottalt lukke.

Akustisk fremhæves knirk eller uregelmæssige svingninger ofte som et træk, men det tillægges ikke samme vægt i alle fremstillinger. Andre træk der fremdrages, men som også tillægges noget varierende vægt er udsving i lydstyrke og grundtone (f_0).

2 2c kunne lige så vel formuleres som en fonotaktisk begrænsning baseret på observeret distribution. For Københavnsk standardsprog kunne den lyde således: Stød forekommer kun i stavelser med lang vokal eller med kort vokal fulgt af sonorant konsonant. Der er en række forhold som taler for at begrænsningerne i distributionen bør betragtes som fonotaktisk (eller fonologisk) snarere end fonetisk betingede:

- Distinktionen +/- stød kan udmærket opretholdes i hvisket tale (som pr. definition udelukker stemthed). Med mindre det antages at stød realiseres på en grundlæggende anden måde i hvisket tale, så harmonerer det dårligt med at der skulle være et fonetisk krav om et vist stræk af stemt lyd for at stød kan realiseres.
- Hurtig tale medfører ikke systematisk tab af stød. Hvis 2c er en *fonetisk* betingelse, kan man undre sig over at stød tilsyneladende ikke svækkes eller tabes ved hurtig tale, hvor man kunne formode at også stavelser med stød ville blive kortere.
- Der gælder ikke de samme distributionelle restriktionerne i alle varianter af dansk. I jysk kan der fx ikke forekomme stød på sonorant konsonant hvis sonoranten følges af endnu en konsonant (inden for samme stavelse) som fx. i ”hals”.

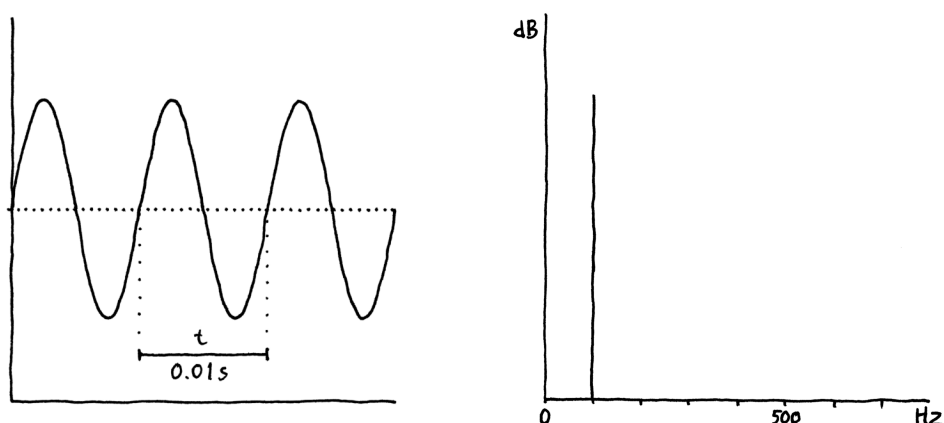
3 Ud over en rent impressionistisk vurdering af tydelighed er det ikke veletableret hvad der kendetegner mere henholdsvis mindre distinkte stød.

For at kunne danne sig et klarere billede af hvordan stød er blevet karakteriseret akustisk og artikulatorisk, følger i kapitel 3 en gennemgang af resultater fra væsentlige nyere fonetiske undersøgelser af stød.

2.2 Nogle akustiske grundbegreber

Det følgende er tænkt som en kort introduktion til nogle af de akustiske grundbegreber der vil optræde i de efterfølgende afsnit. For en mere udførlig ikke matematisk introduktion til akustik kan der henvises til fx Johnson 1997 eller Ladefoged 1962.

Sinusfunktionen danner grundlag for den mest basale type af periodiske signaler. Et kontinuert⁴ sinusformet signal kan udtømmende karakteriseres ved sin frekvens, amplitude (dvs. styrke) og fase (der kan opfattes som det tidsmæssige offset for kurveforløbet). Et kontinuert sinusformet signal udtrykker kun energi ved én bestemt frekvens, som er bestemt af periodelængden. Forholdet mellem periodelængden (t) og frekvensen (f_0) er $f_0=1/t$. Se figur 2.1.



Figur 2.1. Til venstre er amplitude plottet som funktion af tid for et sinusformet signal med periodetiden t på 0.01 sekund. Til højre ses den tilhørende spektralfordeling som illustrerer at der kun er energi til stede ved én frekvens som er $1/t$ dvs. $1/0.01=100\text{ Hz}$.

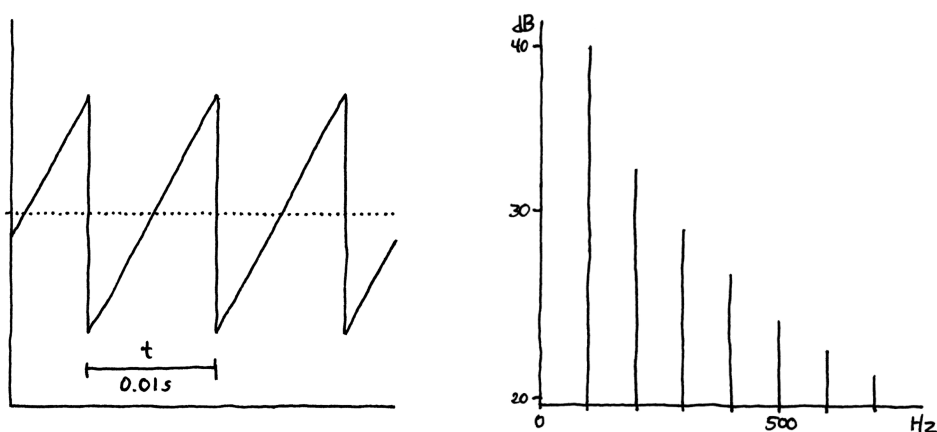
Det kan vises at det ved at kombinere sinusformede signaler er muligt at opbygge (strengt taget approximere) alle komplekse signaler – både periodiske og aperiodiske (se videre nedenfor). Alle signaler kan derfor også opløses i en række af sinustoner karakteriseret ved hver deres frekvens, styrke og faseforskydning. En sådan opløsning af et signal i sinustoner kaldes for fourieranalyse. En opregning af disse sinustoners amplitude⁵ ordnet efter deres frekvens giver tilsammen en beskrivelse af signalets (energi-)spektrum, også kaldet spektralfordeling.

4 Med kontinuert forstås at det er konstant og uændret over tid, og principelt af uendelig varighed.

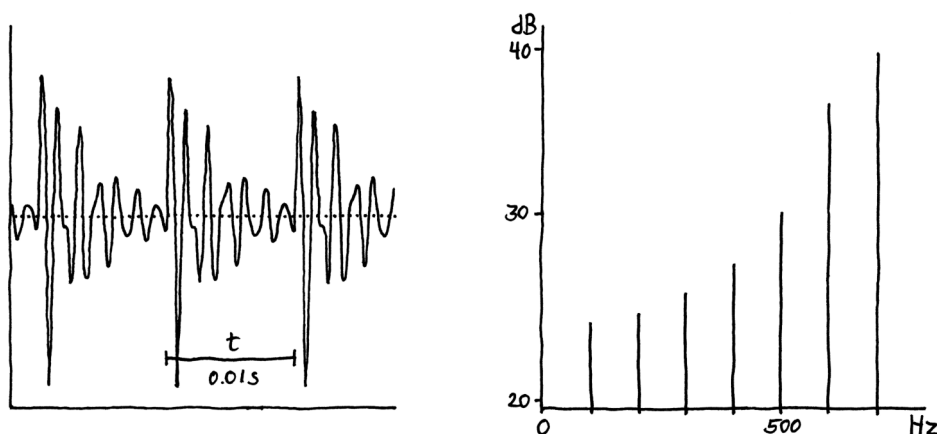
5 Som regel er det deltonernes relative amplitude man interesserer sig for, snarere end de absolutte værdier.

Et sinusformet signal opløses ikke yderligere ved en fouriertransformation (eller rettere, opløses til sig selv), og kan derfor betegnes som et simpelt periodisk signal.

Et kontinuert periodisk signal som ikke følger en sinusform (se eksempel i figur 2.2), vil indeholde energi ved flere frekvenser: Ved grundtonens⁶ frekvens (f_0) som er bestemt af periodelængden, og ved overtoner hvis frekvens er hele multipla af f_0 . Afstanden mellem to nærliggende deltoner er altså altid lig med f_0 . Da et ikke-sinusformet kontinuert periodisk signal indeholder energi ved flere frekvenser, betegnes det som et komplekst periodisk signal.



Figur 2.2. Amplitude som funktion af tid for et trekant-signal med samme periodetid t som i figur 2.1. Den tilhørende spektralfordeling viser at der er energi til stede ved flere frekvenser.



Figur 2.3. Amplitude som funktion af tid og den tilhørende spektralfordeling for et andet komplekst signal med samme periodetid t som i figur 2.1 og 2.2.

⁶ Bemærk at termen "grundtone" har en anden betydning i musikalske sammenhænge.

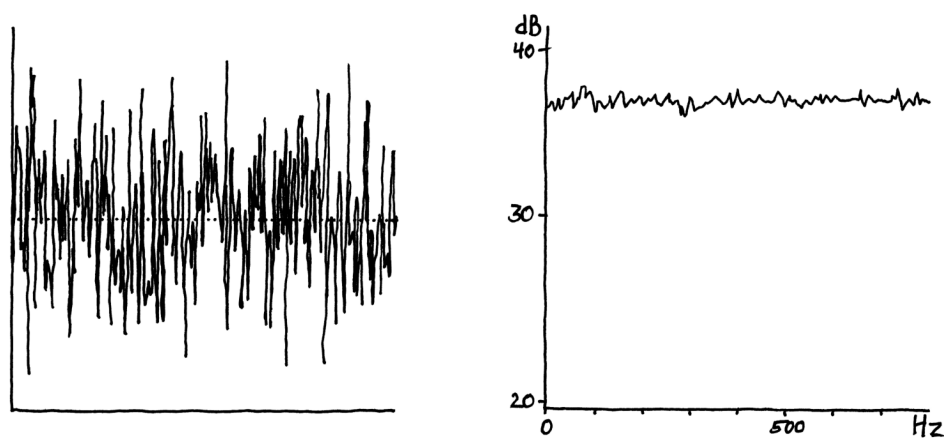
Grundtonen og overtonerne kaldes under et for deltoner. Deltonerne benævnes H1, H2, H3 etc. Den laveste deltone, grundtonen, betegnes således både f_0 og H1. De højere deltoner kaldes også under et for overtoner, og H2 er altså den første overtone. Se figur 2.2

Forskellige komplekse periodiske signaler med samme grundtone vil have deltoner ved de samme frekvenser, men deltonerne vil have forskellige indbyrdes styrkeforhold; dvs. signalernes spektra vil være forskellige, sammenlign figur 2.2 og 2.3. Den hørbare konsekvens er at de vil have samme tonehøjde, men forskellig klang.

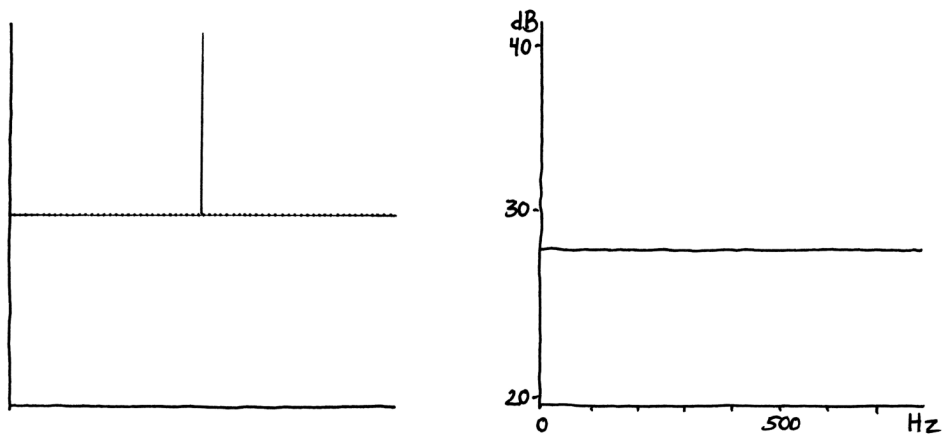
Både simple og komplekse *perfekt periodiske* signaler vil have et diskret spektrum som *kun* indeholder energi ved frekvenser som er hele multipla af grundtonen, jfr. figur 2.1, 2.2 og 2.3.

Aperiodiske signaler har et kontinuert spektrum. Det kan forklares på denne måde: Eftersom et aperiodisk signal netop er defineret ved ikke at gentage sig selv, vil periodelængden være uendelig lang, og dermed bliver grundtonen uendeligt lav. Det medfører at afstanden mellem de enkelte deltoner bliver infinitesimal (uendeligt lille). Dermed bliver spektret kontinuert.

Der er to arketypiske aperiodiske signaler hvis spektrale karakteristika det er værd at hæfte sig ved: Hvid støj er kontinuert aperiodisk signal som indeholder lige meget energi ved alle frekvenser. Se figur 2.4.

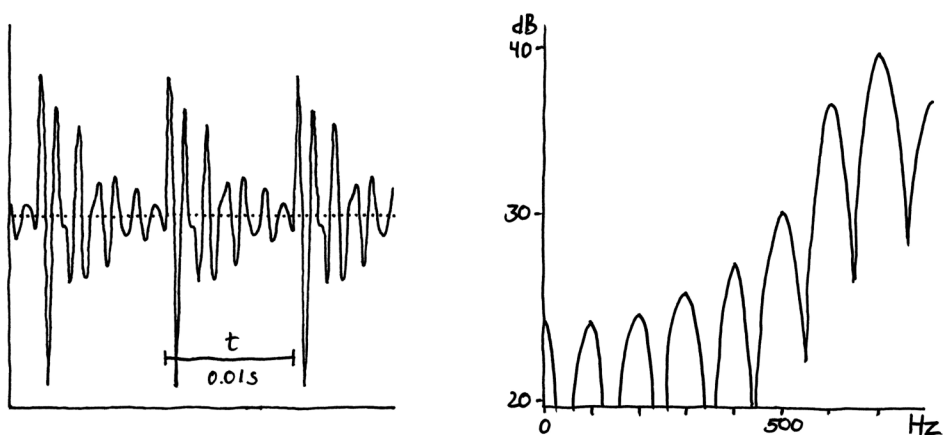


Figur 2.4. Eksempel på amplitude som funktion af tid og den tilhørende spektralfordeling for en sekvens af hvid støj.



Figur 2.5. Amplitude som funktion af tid og den tilhørende spektralfordeling for en impuls.

En uendelig kort impuls vil også indeholde lige meget energi ved alle frekvenser. Se figur 2.5.



Figur 2.6. Amplitude som funktion af tid for et overtonerigt signal (samme som i figur 2.3) og det resulterende kontinuerte spektrum for en spektral analyse med en endelig frekvensopløsning.

De ovenfor nævnte typer af signaler er alle matematiske abstraktioner. Når man undersøger reelt forekommende signaler er man i sagens natur nødt til at se på et endeligt udsnit af det forekommende signal. Selv hvis det signal man undersøger skulle være perfekt periodisk, så vil den del der analyseres kun udgøre en tilnærmelse til et perfekt periodisk signal, fordi tidsudstrækningen er begrænset. Signaler der forekommer at være periodiske inden for det interval der betragtes benævnes quasiperiodiske signaler. Et quasiperiodisk signal vil have spektrale toppe ved deltonefrekvenserne, men det vil have et kontinuert spektrum, sammenlign figur 2.2 og 2.6.

2.3 Modulation: variation over temaet periodiske signaler

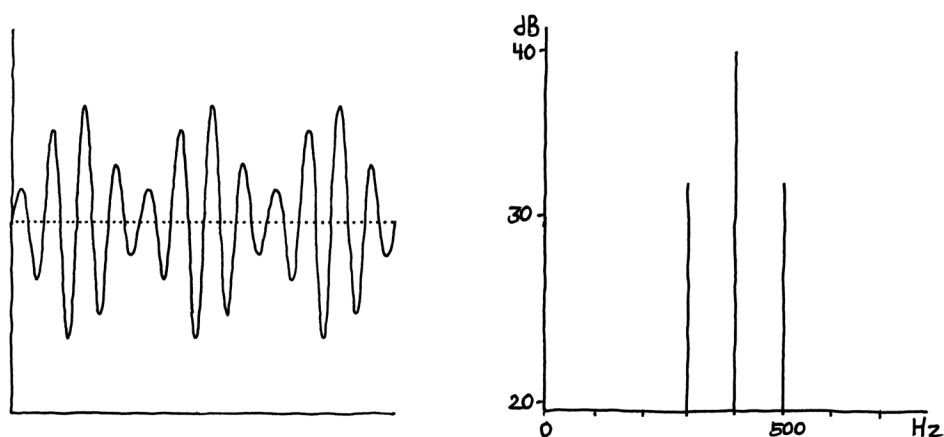
Blant andet i forbindelse med stød kan der forekomme en type variation i stemmelæbernes svingningsmønster som benævnes diplofoni – se videre afsnit 2.4.1. Da der viser sig at være nogle muligheder forbundet med at anskue diplofoniske svingninger som modulation af et periodisk signal, følger her en kort introduktion til begreberne amplitude- og frekvensmodulation.

2.3.1 Amplitudemodulation

Tager man udgangspunkt i en 400 Hz sinustone og lader styrken af tonen variere op og ned i takt med en sinustone på 1 Hz, så vil man have en konstant tone (frekvens) som stiger og falder i styrke én gang i sekundet. Styrken (amplituden) af sinustonen på 400 Hz er nu moduleret med 1 Hz.

Modulationsstyrken bestemmer hvor meget amplituden vil variere op og ned, og modulationsfrekvensen afgør, som nævnt, hvor mange gange i sekundet styrken varierer.

Spektret for en 400 Hz tone amplitudemoduleret med 100 Hz er illustreret i figur 2.7. Der optræder energi ved tre frekvenser: ved grundtonen på 400 Hz og ved to såkaldte sidebånd 100 Hz over og under grundtonen. Havde vi moduleret med 10 Hz, havde sidebåndene ligget 10 Hz over og under grundtonen, som her er 400 Hz.



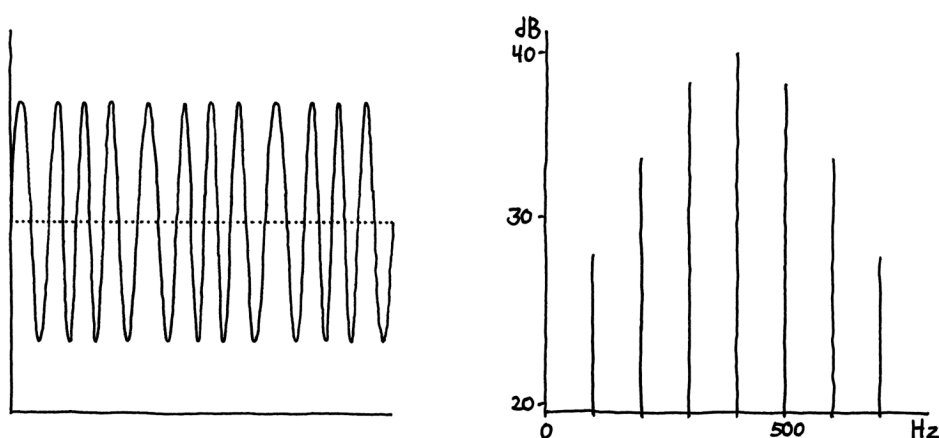
Figur 2.7. Amplitude som funktion af tid og den tilhørende spektralfordeling for en sinustone på 400 Hz amplitudemoduleret med 100 Hz.

Modulerer man et komplekst periodisk signal (dvs. et signal som har overtoner), så vil der tilsvarende dannes sidebånd omkring alle overtonerne. Samlet kan man sige at modulation (ikke overraskende) bidrager til at gøre signalet mere komplekst, og gør frekvensspektret mere energirigt.

2.3.2 Frekvensmodulation

Ligesom styrken kan moduleres, kan man tilsvarende modulere grundtonens frekvens. Hvis man fx tænker sig at lade en 100 Hz sinustones frekvens variere op og ned – fx fra 95 til 105 Hz i takt med en sinuskurve på 1 Hz, dvs. én gang i sekundet – vil tonen være frekvensmoduleret med en 1 Hz sinustone, med en styrke på ± 5 Hz

Spektret for en sinustone på fx 400 Hz frekvensmoduleret med en anden sinustone på fx 100 Hz vil dels indeholde energi ved grundtonen, og dels ved sidebånd over og under grundtonen ved frekvenser som er hele multipla af modulationsfrekvensen, dvs. her $\pm 100, 200, 300$ Hz etc. -se figur 2.8. Sidebåndsfrekvenserne vil overordnet set være svagere jo længere væk fra grundtonen de er, men de falder ikke nødvendigvis jævnt i styrke med afstanden til grundtonen.



Figur 2.8. Amplitude som funktion af tid og den tilhørende spektralfordeling for en sinustone på 400 Hz frekvensmoduleret med 100 Hz.

Frekvensmodulation af et komplekst periodisk signal vil også give sidebånd omkring overtonefrekvenserne, præcis som det er tilfældet ved amplitudemodulation. I de fleste tilfælde vil frekvensmodulation give energi ved flere sidebåndsfrekvenser end amplitudemodulation.

2.3.3 Modulation generelt

Der er i princippet ingen grænser for hvilke signaler der kan moduleres, og hvilke signaler der kan moduleres med. Man kan derfor fx sagtens have en modulationsfrekvens som ligger tæt på, eller over, grundfrekvensen.

Hvis både det modulerede og det modulerende signal er kontinuerte periodiske signaler, vil resultatet også være et kontinuert periodisk signal og dermed et signal der har en diskret række af deltoner. Det resulterende spektrum får dog en mere kompleks sammensætning med flere deltoner end det umodulerede signal havde.

Hvis enten det modulerede eller det modulerende signal er aperiodisk vil det resulterende signal også være aperiodisk, og dermed have et kontinuert spektrum.

Både frekvens- og amplitudemodulation findes naturligt forekommende, og udnyttes teknisk, fx i forbindelse med traditionel radiotransmission – deraf navnene FM og AM-radio. I musikalske sammenhænge forekommer begge typer modulation også på instrumenter som tillader det, fx elektroniske instrumenter (elektriske orgler, synthesizere), strengeinstrumenter og ved sang. Den terminologi som benyttes i musikalske sammenhænge er ikke helt fasttømret, men hurtig modulation af frekvens (bør) benævnes vibrato, mens hurtig modulation af styrke (somme tider opnået ved at anslå den samme tone meget hurtigt efter hinanden) (bør) benævnes tremolo.

Det er måske på sin plads at bemærke at selvom principperne er de samme, så lyder modulation med lave frekvenser under ca. 10-30 Hz væsentligt anderledes end modulation med højere frekvenser. Når modulationsfrekvenserne er lave, opfatter vi den varierende styrke eller tonehøjde som det det er – nemlig variation i styrke eller frekvens. Ved modulationsfrekvenser over ca. 30-50 Hz opfatter vi det alene som en forandring i klangen af det signal som moduleres. I grænseområdet mellem de to kan indtrykket måske bedst beskrives som at det lyder "groft" eller "ruflet".

Grænserne mellem de to opfattelser er i overensstemmelse med hvordan vi i øvrigt opfatter periodiske signaler. Har de en høj frekvens, opfattes de som toner, mens periodiske signaler med lavere frekvenser opfattes som serier af slag (impulser). Grænsen er heller ikke her knivskarp, og er sandsynligvis afhængig af bl.a. signalets karakter (klang).

2.4 Stemmekildens funktionsmåde og akustiske egenskaber

Stemmelæberne fungerer som en effektiv lyd giver ved at dele en kontinuert luftstrøm fra lungerne op i en serie af luftpulser ved at åbne og lukke for luftstrømmen. De trykvariationer som luftpulserne afstedkommer, er det som vi opfatter som lyden af stemmen.

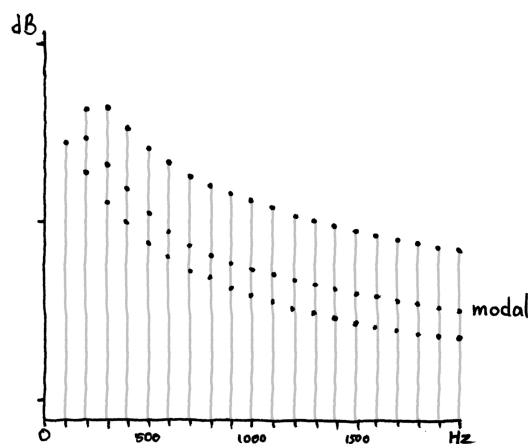
Stemmelæberne sættes i bevægelse ved at lufttrykket fra lungerne tvinger stemmelæberne fra hinanden, men den udstrømmende luft suger hurtigt stemmelæberne mod hinanden igen på grund af den såkaldte Bernoulli-effekt. Når stemmelæberne har lukket for luftstrømmen, opbygges der hurtigt et overtryk under stemmelæberne som igen presser dem fra hinanden, og processen gentager sig.

Der er flere forhold vedrørende de luftpulser der bliver genereret, som er afgørende for hvordan stemmen lyder:

- 1: Størrelsen af luftstrømsforandringen i hver periode bestemmer styrken af den lyd der genereres.
- 2: Det tempo som stemmelæberne svinger med, og dermed tempoet som pulserne følger efter hinanden, bestemmer grundtonen også betegnet f_0 .
- 3: Nøjagtigt hvordan skiftene fra lukket til åben glottis og tilbage igen foregår, har betydning for luftstrømsfunktionens forløb i hver periode dvs. de enkelte pulsers form og dermed lyden fra stemmekilden, dvs. stemmens klang.

Jo mere abrupt åbningen og lukningen af den glottale åbning foregår, jo mere abrupte bliver trykforandringerne hvilket medfører relativt mere energi i lydsignalet ved højere frekvenser dvs. ved højere deltoner. Den hørbare konsekvens vil være en skarpere klang. Når trykforandringerne foregår mere abrupt end normalt⁷, taler man om at stemmen er overkomprimeret eller hyperfunktionel.

Modsat vil mindre abrupt åbning og lukning af glottis give mindre abrupte trykforandringere hvilket medfører mere energi ved grundtonen og relativt mindre energi i lydsignalet ved højere deltoner. Den hørbare effekt er en blødere eller rundere klang. I det tilfælde taler man om at stemmen er underkomprimeret eller hypofunktionel. Figur 2.9 illustrerer hvordan stemmekildens spektrum varierer afhængig af hvor komprimeret stemmen er. Der følger en mere detaljeret gennemgang af emnet i afsnit 5.2.



Figur 2.9. Skitsering af kildespektret for modal samt under- og overkomprimeret stemme.

- 4: Stemmælæbesvingningernes regelmæssighed spiller også en rolle for stemmens klang. Når stemmelæberne genererer pulser i et jævnt ensartet gentaget mønster, er det pulsernes form der bestemmer den spektrale sammensætning af lyden fra stemmekilden og dermed klangen.

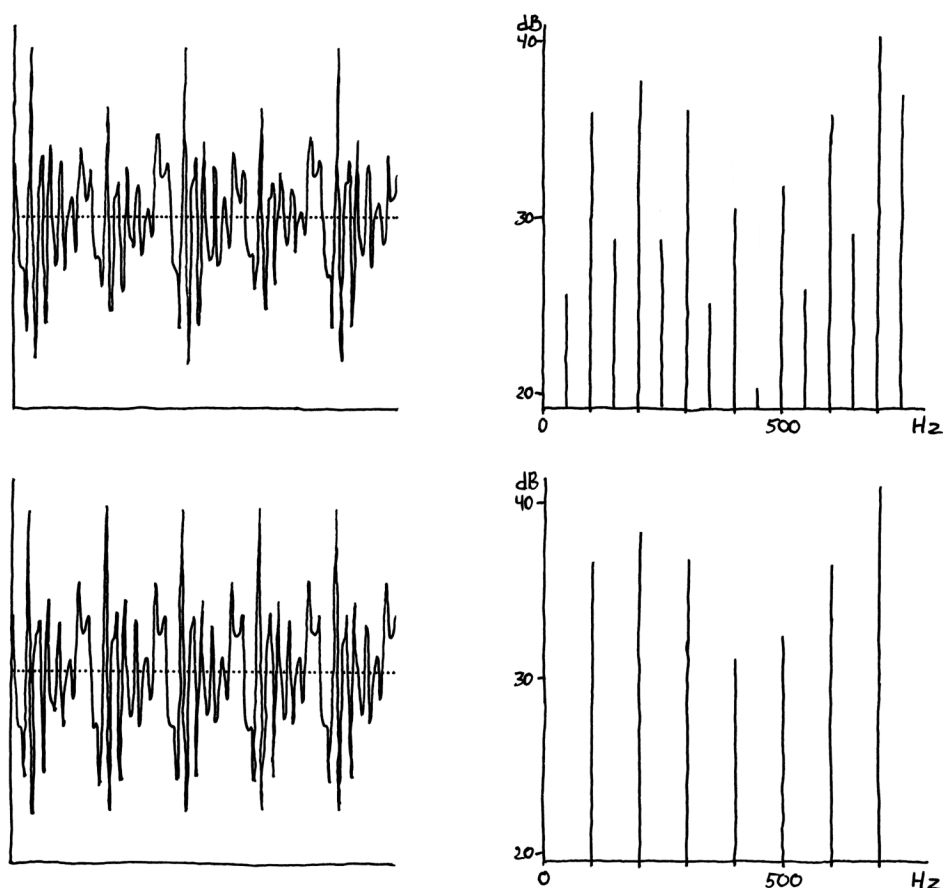
⁷ Se videre afsnit 2.4.3.

Men stemmelæberne kan også åbne og lukke for luftstrømmen i et aperiodisk – kaotisk – mønster. I det tilfælde giver det ikke mening at tale om en tonehøjde (f_0), og lydens spektrum vil heller ikke udvise en struktur med deltoner. Det klangmæssige resultat kan beskrives som groft eller raspende. Se videre afsnit 5.5 og frem.

Imellem disse to yderpunkter findes et kontinuum af mere eller mindre jævn periodicitet. Man taler om jitter som mål for ustabilitet i timingen af de glottale pulser (og dermed ustabilitet i grundtonens frekvens), og om shimmer som det tilsvarende mål for ustabilitet i de glottale pulsers amplitude.

2.4.1 Mellem modal stemthed og knirk: diplofoni

Variation i periodiciteten behøver ikke være tilfældig – den kan også være systematisk. Der optræder ikke sjældent hvad man kunne kalde for dobbeltsvingninger i form af en jævn vekslen mellem større og mindre pulser, og/eller mellem perioder med længere og kortere varighed sådan at hveranden periode er kraftigere og/eller længere. Når stemmekilden følger et sådant svingningsmønster, betegnes det diplofoni.



Figur 2.10. Amplitude som funktion af tid og spektrum for et diplofonisk signal (øverst) og figurer for det tilsvarende signal uden diplofoni (nederst).

Akustisk set kan variation i periodiciteten anskues som modulation. Betragter man diplofoni som modulation, er det et specialtilfælde hvor modulationsfrekvensen er den halve af grundtonen. I forhold til spektret har det den konsekvens at de øvre og nedre sidebånd falder sammen idet deltonerne jo ligger med en afstand svarende til grundtonens frekvens, og sidebåndene ligger i en afstand over og under deltonerne der er bestemt af modulationsfrekvensen (som her er den halve af grundtonens frekvens), se figur 2.10. Resultatet er at spektret kommer til at se ud som om det stammede fra et signal med den halve grundtone, og man kan naturligvis også godt kan argumentere for at grundtonen nu faktisk er den halve (dvs. lig med modulationsfrekvensen).

Der er dog gode grunde til at betragte diplofoni som et modulationsfænomen: Ved almindelig modal stemthed modsvarer en periode i det akustiske signal et periodisk tilbagevendende fysisk forløb, nemlig at stemmelæberne åbner og lukker for luftstrømmen. Hvis den diplofoniske stemme skulle analyseres som et signal med den halve grundtone (svarende til modulationsfrekvensen), ville det have som konsekvens at en periode i det akustiske signal kom til at svare til en sekvens af to åbne- og lukkefaser. Ved at anskue det som modulation kan man bibeholde den enkle sammenhæng mellem det akustiske signal og de fysiologiske hændelser. Det stemmer også overens med at der skal en betydelig grad af modulation til før det påvirker opfattelsen af stemmens tonehøjde.

2.4.2 Glottis som støjkilde

Der kan også dannes lyd i glottis (og andre steder i luftkanalen) ved at presse luft igennem en passende snæver passage sådan at der opstår turbulens i luftstrømmen. Det resulterer i en støjende lyd. I sin mest udprægede form kender vi det som hvissen, hvor stemmelyden erstattes af den turbulente støj. Der kan også forekomme en større eller mindre andel af turbulent støj samtidig med at stemmelæberne svinger. Støjen fremkommer ved at stemmelæberne ikke lukker fuldstændig af for luftstrømmen under lukningen, og dette giver anledning til det man kalder luftfyldt stemme.

2.4.3 Indkredsning af begrebet modal stemthed

Modal stemthed er den "almindelige" upåfaldende måde at bruge stemmen på. Almindelige modal stemme er karakteriseret ved medium kompression, ved at have en jævn periodicitet og ved ikke at være luftfyldt. Modal stemme er dermed i høj grad negativt defineret ved *ikke* at være luftfyldt, ved *ikke* at være underkomprimeret, ved *ikke* at være overkomprimeret, og ved *ikke* at have uregelmæssigheder i svingningsmønsteret. Definitionen slipper dog ikke for et element af cirkularitet da over- og underkomprimeret kun giver mening med reference til den (normale) modale grad af kompression. Inden for logopædien beskrives "modal" som den sundeste og mest effektive måde at bruge stemmen på. Det betyder dog ikke at det er den måde at bruge stemmen på der resulterer i den kraftigste lyd hverken akustisk eller auditivt.

3 Baggrund: tidligere undersøgelser af stød

Stød som det forekommer i Dansk rigsmål (se afsnit 2.1). har været genstand for en række instrumentalfonetiske studier i løbet af det tyvende århundrede. Smith 1944, Riber Petersen 1973, Thorsen 1974, Fischer-Jørgensen 1987, 1989a, 1989b, samt en række artikler af Grønnum & Basbøll fra 2001 til 2012, må betragtes som de væsentligste moderne instrumentelt funderede kilder til viden om stødets fonetik.

Der er enkelte instrumentalfonetiske undersøgelser af stød i andre varianter af dansk – først og fremmest Ringgaard 1960, der omhandler vestjysk stød, og Køst 2008, der omhandler stød i Århusiansk regiolekt. Stød som det realiseres i forskellige varianter af dansk er et emne der fortjener større opmærksomhed(!) men som jeg ikke vil få anledning til at komme nærmere ind på i denne undersøgelse, der også tager udgangspunkt i stød som det forekommer i Dansk rigsmål.

Smith 1944 kan betegnes som den første moderne instrumentelle undersøgelse af stød i den forstand at der (til dels) benyttes elektriske metoder til registrering af lyd og muskelaktivitet, hvor tidligere instrumentalfonetiske undersøgelser af stød har været henvist til at benytte sig af rent mekaniske metoder.

Svend Smith undersøgte en forklaringsmodel der gik på at stød primært skulle skyldes en kontraktion i de ekspiratoriske muskler, som på grund af trykforøgelsen skulle afstedkomme en reflektorisk modreaktion i form af en kontraktion i glottis.

Undersøgelsen bygger på registrering af EMG-aktivitet i de abdominale muskler, luftstrøm (ved munden) og registrering af akustisk målbare forhold som intensitet, grundtone (f_0) og forekomst af uregelmæssige svingninger (herunder diplofoni).

På grund af vanskeligheder med at holde de forskellige registreringer tilstrækkeligt præcist synkroniserede, var der ikke muligt at afklare om forklaringsmodellen holdt.

Riber Petersen 1973 er en rent akustiske undersøgelse af stød, som ser på varighed, intensitet, grundtone (f_0), og forekomst af uregelmæssige svingninger.

Riber Petersen fremhæver at der ses en påfaldende variabilitet i stødets akustiske fremtoning. Navnlig en enkelt informants stød påkalder sig opmærksomhed ved ikke at give sig tydeligt til kende i nogen af de akustiske analyser, på trods af at disse stød auditivt forekommer lige så tydelige som de øvrige informanternes stød.

Thorsen 1974 er et pilot-perceptionsforsøg baseret på stød fra den informant der skilte sig ud i Riber Petersens undersøgelse.

Perceptionsforsøget viste at lyttere med større sikkerhed kunne identificere stød vs. ikke stød i ordpar fra denne taler, sammenlignet med ordpar udtalt af en taler der producerede stød som udviste tydelige og prototypiske akustiske karakteristika. I forlængelse af det overraskende resultat, gennemførtes et forsøg hvor succesivt kortere dele af to ordpar med og uden stød blev præsenteret for lyttere. Ideen var at finde ud af om der var forskel på de prototypiske og de ikke-synlige stød med hensyn til hvor tidligt i ordet stødet kunne opfattes. Resultatet var at stødet i den prototypiske form ikke blev opfattet når der blev klippet før de uregelmæssige svingninger satte ind, hvorimod de 'usynlige' stød blev opfattet selv når der blev klippet kort efter vokalens begyndelse. Det kunne tyde på at der i det ikke synlige stød fandtes hørbare tegn på stød noget tidligere i stavelsen.

Imidlertid adskilte ordet med det ikke-synlige stød sig også fra de øvrige ord ved et lidt anderledes grundtoneforløb, hvilket kom til at stå i vejen for en klar tolkning af resultatet, eftersom det ikke kunne udelukkes at lytterne havde brugt dette træk til at identificere ordet med det ikke-synlige stød.

Fischer-Jørgensen, 1989a (samt 1987 og 1989b)⁸, udgør en metodisk meget vidt favnende akustisk og fysiologisk fonetisk undersøgelse af stød der kombinerer mange forskellige instrumentelle tilgangsvinkler og som trækker på bidrag fra en række medarbejdere og samarbejdspartnere. Afsnit 3.1 er viet til en mere detaljeret gennemgang af resultaterne og et par af de centrale diskussioner og tolkninger.

Foranlediget af Basbølls morabaserede analyse af stød og med afsæt i Fischer-Jørgensen 1987, 1989a og 1989b har Grønnum & Basbøll i samarbejde stået for en række undersøgelser, beskrevet i en række artikler fra 2001 til 2007. Disse artikler vil blive gennengået i afsnit 3.2.

Med hensyn til øvrige tidligere undersøgelser af stød henvises til gennemgangen i Fischer-Jørgensen 1989a.

Når man sammenligner de forskellige undersøgelser af stød er der slående forskel på den vægt forskellige aspekter ved stødrealisationen tillægges når de primære træk ved stød gøres op. Som afslutning på kapitlet følger i afsnit 3.3 en diskussion centreret omkring en af de mere påfaldende forskelle.

8 Fischer-Jørgensen 1989a er en let revideret genudgivelse af 1987. Fischer-Jørgensen 1989b er en kortere artikel baseret på de samme undersøgelser.

3.1 Fischer-Jørgensen

Undersøgelserne i Fischer-Jørgensen 1989a falder inden for den tradition der betegnes laboratoriefonetik. Det sproglige materiale undersøgelserne bygger på, er oplæst tale baseret på lister af enkelt-ord eller ord placeret i såkaldte bæresætninger⁹.

De akustiske analyser omfatter måling af varighed, intensitet og f_0 (grundtone-forløb), forekomst af uregelmæssige svingninger og evt. glottalt lukke, spektrale analyser (NB- og WB-spektrogrammer og spektralplots) samt estimering af kildesignal fra glottis og af luftstrøm gennem glottis ved hjælp af invers filtrering.

De fysiologiske analyser omfatter elektromyografisk registrering af muskelaktivitet (EMG) i bl.a. vocalis og andre muskler i larynx, måling af lufttryk over og under glottis, måling af luftstrøm gennem talekanalen, palatogrammer, registrering af larynxposition og videooptagelse af stemmelæberne ved hjælp af fiberskop.

Fischer-Jørgensen 1989a betragter, med henvisning til Svend Smith 1944, stavelser med stød som delt i to faser. Fischer-Jørgensen betegner den anden fase som den egentlige stødfase¹⁰. Det beskrives sådan at de to faser groft taget falder i henholdsvis første og anden halvdel af en lang vokal med stød. Når der er tale om stød på kort vokal efterfulgt af sonorant konsonant, falder de to faser stort set sammen med henholdsvis vokalen og konsonanten. Fischer-Jørgensen påpeger dog (p. 128) at der ikke nødvendigvis er en klar tidsmæssig afgrænsning af faserne. Fx kan der være en mere eller mindre glidende overgang mellem de to faser, og de kan strække sig ind over nabosegmenter. Inddelingen af stød i to faser må derfor opfattes som en prototype-ide. De akustiske og fysiologiske forhold er dog (oftest) markant forskellige i de to faser. Derfor opsummeres resultaterne for hver fase for sig.

Nedenstående tabeller 3.1 og 3.2 er beregnet på at give et overblik over resultaterne som gennemgås detaljeret i de efterfølgende afsnit. Der er meget stor forskel på hvor omfattende materiale der ligger bag de forskellige undersøgelser. Derfor angives det i gennemgangen hvor mange ord eller ordpar og hvor mange informanter de enkelte delundersøgelser er baseret på. Efter gennemgangen følger en kort oversigt over resultaterne i afsnit 3.1.3, der med fordel kan læses sammen med de to nævnte tabeller.

9 Et eksempel på en bæresætning kunne være "Det er _____ de siger".

10 Heri afviger Fischer-Jørgensen fra Smith der betegner den første fase som stødfasen, og den anden fase som en reflektorisk modreaktion i glottis.

Akustiske resultater			
Betonede stavelse			Ubetonet stavelse
	Fase 1	Fase 2	
Konsonant	Lang vokal		
Konsonant	Kort vokal	Sonorant konsonant	
	Lange vokaler med stød er kortere end lange vokaler uden stød, men kun i diskontinuerlig tale		
	Kort vokal før konsonant med stød er kortere end tilsvarende vokaler før konsonant uden stød	Konsonanter med stød er længere end tilsvarende konsonanter uden stød	
	f_0 altid lidt højere sammenlignet med ord uden stød	Fald i f_0 efterfulgt af eller kombineret med aperiodicitet 53.8 %	f_0 (lidt) lavere sammenlignet med ord uden stød
	f_0 : ord med stød kan have høj-lav tonalt forløb, modsat normalt lav-høj-forløb.		
		Uregelmæssige svingninger (aperiodicitet) 70.8 %	
	Intensitet ofte lidt højere sammenlignet med ord uden stød	Kraftigt fald i intensitet	
		Ikke glottalt lukke	
	Relativt mere energi ved højere formanter (F4 og F5)	Fald i styrke for de lavere deltoner	
		Invers filtrering tyder på mindskede neg. spikes	
		Invers filtrering tyder på faldende glottal luftstrøm og længere lukkefase	

Tabel 3.1. Oversigt over akustiske resultater fra Fischer-Jørgensen 1989a.

Fysiologiske resultater		
Betonet stavelse		
Konsonant	Fase 1	Fase 2
Kraftigere aktivitet i PCA i prævokaliske ustemte konsonanter	Snævrere artikulation af vokalen [i] i stavelser med stød ifht. stavelser uden stød	
	Aktivitet i vocalismusklen som begynder i vokalens første fase, men som når sit maksimum kort efter det tidspunkt hvor faldet i grundtonen eller uregelmæssighederne begynder. Tilsvarende peak ses for LCA.	
	Højere lufttryk under glottis	Fald i trykket under glottis (dog stadig højere end i ord uden stød)
	Højere lufttryk i esophagous (ud for glottis)	
		Konstriktion af stemmelæberne (vocalis) Adduktion af de falske stemmelæber (ventricular folds)
Lufttrykket i farynx (over glottis) muligvis lidt højere i den prævokaliske konsonant		Tendens til lavere tryk i farynx (over glottis)
Kraftigere luftstrøm i prævokalisk konsonant og stejlere stigning		Lavere luftstrøm gennem munden og dermed også glottis

Tabel 3.2. Oversigt over fysiologiske resultater fra Fischer-Jørgensen 1989a.

3.1.1 Akustiske resultater

Varighed *Vokaler: n=3112 (1556 ordpar), 15 informanter.*

Konsonanter: n=678 (339 ordpar), 7 informanter.

Fischer-Jørgensen (p. 48) finder at lange vokaler med stød oftest er kortere end modsvarende lange vokaler uden stød. Fischer-Jørgensen finder også eksempler hvor der ikke er forskel, eller tilfælde hvor vokalen med stød er længere. Forholdet ser ud til at variere med vokalkvalitet og konsonant-kontekst. Forskellen i varighed opretholdes tilsyneladende kun i distinkt tale. Undersøgelserne der vedrører varighed af lange vokaler med og uden stød bygger på ca. 3112 tokens (1556 ordpar) fra 15 informanter.

I stavelser med stød på konsonant finder Fischer-Jørgensen at konsonanten har længere varighed og vokalen kortere varighed sammenlignet med modsvarende ord uden stød. Undersøgelserne der vedrører varighed af konsonanter med og uden stød, bygger på ca. 678 tokens (339 ordpar) fra 7 informanter.

Fischer-Jørgensen bemærker at observationerne i det store og hele er i overensstemmelse med hvad tidligere undersøgelser har vist, bl.a. Abrahams 1943, Fischer-Jørgensen 1955, Lauritzen (upubliceret) og til dels Riber Petersen 1973.

f_0 generelt *n= ca. 3400 (1700 ordpar), 13 informanter.*

Fischer-Jørgensen opdeler fremlæggelsen af resultaterne der har med grundtone at gøre i tre dele: Resultater der vedrører første del af stavelsen (med stødets første fase), resultater der vedrører anden del af stavelsen (med stødets anden fase, også betegnet som den egentlige stødfase), og endelig resultater der vedrører trykgruppemønsteret (dog her begrænset til den betonedede stavelse og den første efterfølgende ubetonede stavelse). Undersøgelserne der vedrører f_0 , bygger på ca. 3400 tokens (1700 ordpar), fra ialt 13 informanter.

f_0 – første fase

I den første del af stavelser med stød ses altid en lidt højere grundtone (3,5-30 Hz) end i matchede ord uden stød. Forskellene er lidt mindre i enstavelsesord. Forskellen i f_0 i den første del af stavelsen er mere konsistent end den tilsvarende observerede forskel i intensitet (se efterfølgende side).

f_0 – anden fase

Det er værd indledningsvis at knytte en kommentar til de resultater der vedrører f_0 og uregelmæssige svingninger i fase 2: Bestemmelse af grundtoneforløbet foregik ved hjælp af et såkaldt trans pitchmeter¹¹. Det fremgår ikke klart hvor nøjagtigt apparatet var eller hvor robust det var over for uregelmæssigheder i periodicitet, men der er ikke grund til at tro at det har fungeret bedre end de computerbaserede metoder vi har

¹¹ Trans pitchmeteret var et analogt elektronisk måleapparat konstrueret og bygget af Børge Frøkjær-Jensen (på daværende tidspunkt ansat ved fonetisk institut under Fischer-Jørgensen.) Se videre Frøkjær-Jensen 1960 Le Trans Pitchmeter.

til rådighed i dag, og som netop viser deres svagheder jo mindre klar periodiciteten i signalet er. Se videre diskussion af denne problematik i afsnit 5.5. I passager med mere eller mindre uregelmæssige svingninger (som ofte forekommer i fase 2) er det derfor tvivlsomt om systemet fuldt pålideligt har kunnet registrere grundtonen, eller med sikkerhed afgøre om signalet var tilstrækkelig periodisk til at det giver mening at tale om en grundtone eller ej. De resultater der vedrører grundtoneforløbet i stødets anden fase (og forekomsten af uregelmæssige svingninger), må derfor tages med forbehold.

I fase 2 ses i 53.8 % af tokens med stød et fald i grundtone efterfulgt af eller kombineret med uregelmæssige svingninger. Om forekomsten af f_0 -dyk siger Fischer-Jørgensen p. 58 "Irregularity is a characteristic of a strong stød, Fo dip of a weak stød as pronounced by the same speaker"

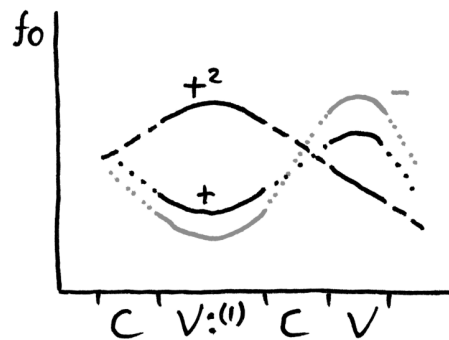
I de tilfælde hvor der ikke registreres et dyk i f_0 , kan der enten være tale om at der er uregelmæssige svingninger og dermed ikke er nogen grundtone, eller at der er jævn eller stigende grundtone.

Den sidste type er sjælden for de fleste informanternes vedkommende. I 20 ud af 26 informanternes sessioner ses denne type i 0-5 % af tokens. For 5 sessioner ligger andelen mellem 7 % og 23 %, mens det for én informant i én optagelse gælder i 62 % af tokens, men samme taler præsterer 0 % af denne type i en anden session.

For visse, men ikke alle, informanter ser fordelingen af denne type ud til at hænge sammen med styrken af stødet: p. 64 "[...] for NR and OB lack of both Fo drop and irregularity is only found in some examples with very weak stød, [...] JR and BRP [...] show a real deviation from the general pattern by having rising Fo in the stressed syllable in many cases of quite distinct stød"

f_0 – trykgruppemønstret

p. 65 "In words with stød the first part of the stressed vowel is higher than in words without stød [...] the following unstressed syllable is generally said on the same pitch level as in words without stød, or sometimes on a slightly lower pitch. The difference between the two syllables is therefore always smaller than in words without stød and in some cases the two syllables are on the same level, or the stressed syllable may even be higher than the following unstressed syllable [...]". Fischer-Jørgensen omtaler den sidste mulighed som "reversed stress contour". Det almindelige forløb for ord uden stød og de to omtalte forløb for ord med stød er illustreret i figur 3.1.



Figur 3.1. Illustration af det tonale forløb dvs. trykgruppemønsteret i ord med og uden stød. Forløbet for ord med stød (markeret med et plus) ligger lidt højere i den betonedede første stavelse, og lidt lavere i den efterfølgende ubetonede stavelse sammenlignet med ord uden stød (grå). I ord med stød ses undertiden et modsat tonalt forløb (markeret med +²), som har sit højeste niveau i den betonedede stavelse.

Informanterne kan inddeles i tre grupper afhængig af hvordan forekomsten af det omvendte (høj-lav) trykgruppemønster er fordelt:

- I Stort set kun i ord med snævre vokaler
 - Ia Få eksempler, hyppigst i ordet [pi:'ber] (HU, NR, ND og PD)
 - Ib Mere almindeligt forekommende med alle typer af snævre vokaler, inklusive [o:] men tilsyneladende ikke [e:]. (BH, BM og FJ)
- II Omvendt trykgruppemønster mulig uanset vokalkvalitet. (JR, PM, OB og BF)

Fischer-Jørgensen bemærker at informanterne¹² der tilhører gruppe Ia, alle er fra København eller Sjælland, mens informanterne i gruppe II har jysk eller fynsk baggrund (hhv. 3 og 1 person). To informanter (BRP, JJ) er tilsyneladende ikke kategoriseret. Det er muligt at de to tilhører en fjerde gruppe som ikke udviser nogen eksempler på omvendt trykgruppemønster. Ellers er det påfaldende at alle de øvrige 11 informanter i et eller andet omfang præsterer dette omvendte trykgruppemønster. Der er altså ikke tale om et eksotisk og sjældent fænomen.

Uregelmæssige svingninger $n = ca. 3400$ (1700 ordpar), 13 informanter.

Undersøgelserne der vedrører uregelmæssige svingninger, bygger på samme materiale som undersøgelserne af f_0 .

Fischer-Jørgensen (p. 50) finder i gennemsnit uregelmæssige svingninger i 70.8 % af tokens med med stød ($n=1700$). Der er dog stor variation fra session til session: fra 8-100 % forekomst. Det gælder også for forskellige optage-sessioner med samme taler.

Timing af uregelmæssige svingninger opsummeres således: "[...] The irregularities (or dip of the F_0 - or intensity curves) are generally found about [100-150 ms] after

¹² Informanterne fremgår p. 182 i Fischer-Jørgensen 1989a.

the start of the vowel; therefore, if the vowel is long they will coincide with the end of the vowel, and if the vowel is short with the beginning of a following sonorant consonant"

Det ser ud til at der oftere er uregelmæssige svingninger i snævre vokaler og i konsonanter. "There are also individual differences [...] but the largest differences are found between different readings by the same speaker. [...] It seems mainly to be a question of the distinctness level chosen by the speaker in the session in question"

Der forekommer ikke knirk i betonede vokaler uden stød, men det forekommer i ubetonede stavelser uden stød.

Forekomsten af uregelmæssige svingninger er i knapt halvdelen af det optagne materiale afgjort ud fra manuelle bedømmelser af duplex oscillogrammer med f_0 og intensitetskurve(r) eller luftstrømskurve (se eksempel p. 51 i Fischer-Jørgensen 1989a). For resten af materialet er klassifikationen alene foretaget ud fra om der forekommer et hul i det f_0 -forløb som trans pitchmeteret har registreret. Efter min opfattelse må de klassifikationer der er foretaget på baggrund af pitchtrackingen, tages med et vist forbehold. Se videre under afsnittet om f_0 , foregående side.

Tegn på forekomst af glottalt lukke $n = \text{ca. } 1700$, 13 informanter.

Fischer-Jørgensen finder ingen klare eksempler på glottalt lukke. Der er dog "[...] some cases where the intensity curve touched the zero-line for a cs, [10 ms] or where the intensity was so low that it might perhaps be perceived as a pause" ($n = \text{ca. } 1700$ ord med stød).

Intensitet $n = 1368-2700$ (684-1350 ordpar), 12 informanter.

Antallet af tokens der ligger til grund for analyserne af intensitet, fremgår ikke helt klart, men må være mellem 1368 (684 ordpar) og ca. 2162-2700 (1081-1350 ordpar) se henholdsvis p. 53 øverst og p. 175. Antallet af informanter er 12.

Intensitetsforløbet for stavelser med og uden stød adskiller sig "in a very consistent way"¹³ (p. 53): Der er tendens til at der nås et lidt højere intensitetsniveau i begyndelsen af vokalen i stavelser med stød. I stødets anden fase forekommer et kraftigt fald i intensitet sådan at der nås et lavere niveau ved vokalens slutning (eller i konsonanten når der er tale om kort vokal med stød i efterfølgende sonorant) sammenlignet med modsvarende ord uden stød. For en enkelt informant (HU) starter faldet i intensitet allerede i den første fase.

13 Det bemærkes dog at tokens fra ét ud af de 110 analyserede typer af ordpar adskiller sig fra det generelle mønster.

Spektrale analyser *n= 112, 14 informanter.*

De spektrale analyser er baseret på et mere beskedent materiale. I alt 112 tokens (56 ordpar) er analyseret. For 44 ordpar er der genereret WB- og NB-spektrogrammer, samt powerspektrum¹⁴. For 12 ordpar er der kun genereret WB-spektrogrammer.

Grunden til at de spektrale analyser er af mere beskedent omfang sammenlignet med de øvrige akustiske analyser, er efter al sandsynlighed at det har været betydeligt mere resourcekrævende at lave spektrale analyser med de metoder der var til rådighed. Hvor visse akustiske analyser som fx mål af intensitet eller f_0 kunne skrives ud på papir med kurveskrivere og dermed kunne genereres i "endeløse baner" i real tid, kunne spektrografen kun behandle få sekunders lydoptagelse af gangen, og krævede adskillige minutter for at skabe hvert spektrogram.

Fischer-Jørgensen ser dels på styrken af højere formanter (F4 og F5) og dels på styrkeforholdet mellem f_0 og F1: (p. 70) "Since the Danish stød has often been described as creaky voice, words with stød may be expected to have relatively strong higher overtones and weak Fo" -se videre afsnit 3.1.4.

Fischer-Jørgensen finder højere intensitet ved F4 og F5 i vokalens første del for 5 ud af 11 talere, men kun sjældent for de øvrige 6 talere.

Med hensyn til forholdet mellem f_0 og F1: (p. 72) "A clear difference [in relative amplitude of F1 and Fo] was found at the end of 21 out of the 25 pairs with open vowels". Én taler producerer dog to tokens hvor det modsatte forhold gør sig gældende. De forskelle Fischer-Jørgensen registrerer, er beskedne (2.2-8.5 dB) sammenlignet med en undersøgelse af Kirk et al. som Fischer-Jørgensen refererer til. I den ses forskelle mellem modal stemthed og knirkestemme på over 10 dB.

"All speakers [except one] have, however, a clear decrease in amplitude in lower harmonics [...] at the end of vowels with stød, whereas in most cases the higher harmonics (particularly around 3-5 kHz) remain unchanged"

Fischer-Jørgensen finder altså tegn på spektrale forskelle mellem stavelser med og uden stød, men ikke så store og ikke så konsekvent forekommende som forventet ud fra de forskelle der tidligere var fundet for fx modal stemthed overfor knirkestemme.

Estimering af kildesignal og glottal flow *n<35, 2 informanter.*

Fischer-Jørgensen præsenterer et pilotforsøg med invers filtrering som giver et estimat af kildesignalet (dvs. lyden fra stemmekilden før den bliver farvet af talekanalens resonanser) og integrering af det estimerede kildesignal som giver et bud på luftstrømsfunktionen.

Da metoden har krævet særligt udstyr og ekspertise, er optagelserne og analyserne udført på KTH i Stockholm i samarbejde med Inger Carlsson, der specialiserede sig i

14 I Praat bruges betegnelsen "spectral slice".

denne metode. Det undersøgte materiale er beskedent. Der er (så vidt man kan se af teksten) tale om mellem 9 og 32 tokens for én taler og tre for en anden.

For den ene talers vedkommende tyder de integrerede kurver (som giver et estimat af glottis luftstrømsfunktion) på en mindsket luftstrøm gennem glottis i den sidste halvdel af vokaler med stød og samtidig en relativt længere lukkefase sammenlignet med ord uden stød og sammenlignet med den første halvdel af vokalen med stød. For den anden talers vedkommende ses kun en beskedent forskel mellem vokal med og uden stød.

Derudover udviser det inverst filtrerede signal mindskede "negative spikes" i stødfasen som angiveligt kan kobles med mindsket overall intensitet i lydsignalet. Se Fischer-Jørgensen 1989a, p. 75 og p. 80.

3.1.2 Fysiologiske resultater

Fiberskop-optagelser *n= ca. 86 (35 ordpar + nogle enkeltord), 7 informanter.*

Fischer-Jørgensen har analyseret ændringer i glottis' konfiguration ved hjælp af videooptagelser, filmet gennem fiberoptisk endoskop (aka fiberskop) indsat via næsen. Da fiberskopet har udsigt til glottis fra en position lige bag drøbelen (som indikeret i figur 3.2), er der kun frit udsyn til stemmelæberne i forbindelse med visse sproglyde. Fx vil tungen dække for udsynet i forbindelse med r-lyde eller bagtungevokaler hvor passagen bag om tungen er snæver. Figur 3.3 viser et billede af glottis fra omtrent denne vinkel. Den anvendte frekvens på 50 billeder i sekundet giver ikke mulighed for at analysere forløbet af de enkelte perioder af stemmelæbernes svingninger (det ville kræve en billedfrekvens som er mindst 2-3 gange så høj som grundtonen), men kan vise ændringer i glottisk konfiguration.

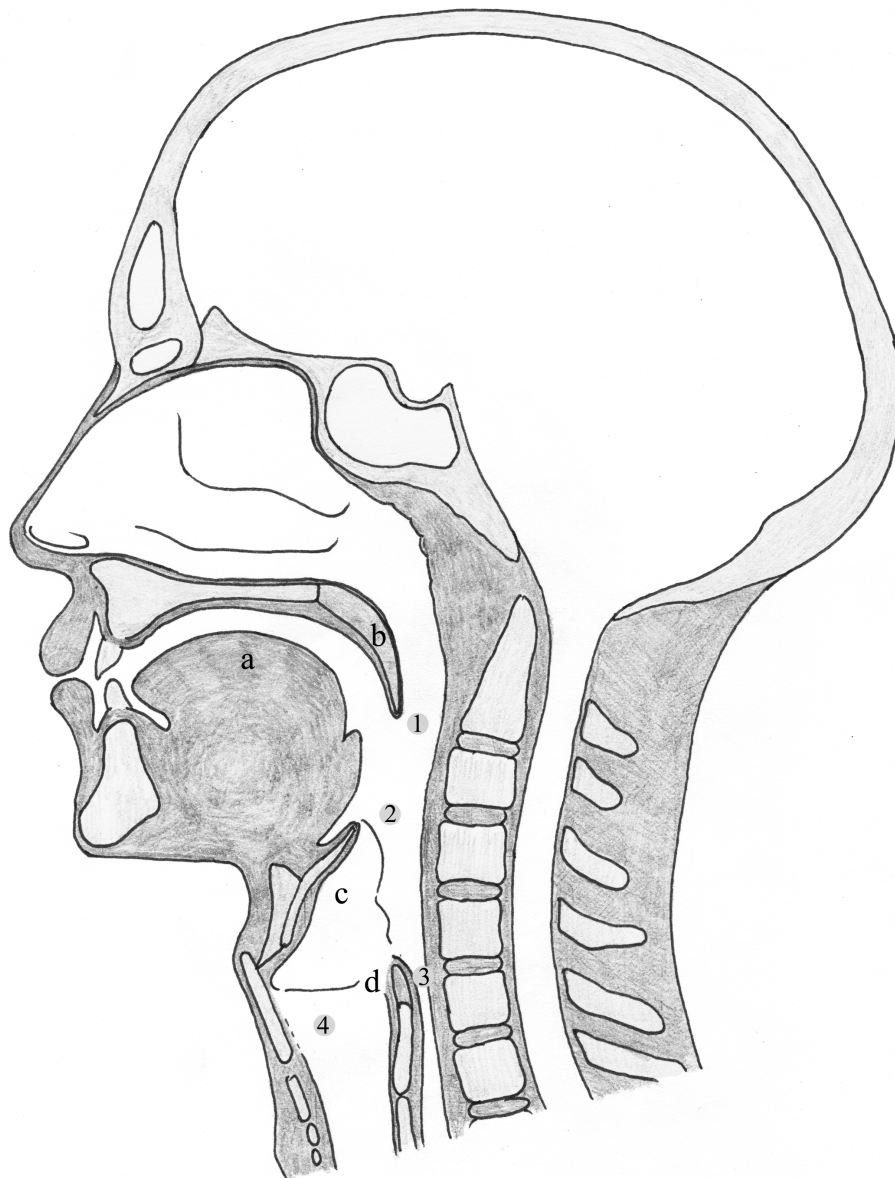
Optagelserne viser i mange tilfælde konstriktion af stemmelæberne (vocalis) og adduktion af de falske stemmelæber (ventricular folds). Der er dog en betydelig variation fra taler til taler.

Lufttryk (under, over og ved glottis)

Lufttrykket er målt under glottis og over glottis, dels i esophagous (dvs. spiserøret ud for glottis) og dels i farynx, se figur 3.2. Lufttrykket over glottis kan måles med en sonde der indføres gennem næsen. Lufttrykket under glottis måles ved at indføre en sonde gennem et hul udefra ind i luftrøret under glottis. Dette indgreb er (af gode grunde) kun udført på en enkelt forsøgsperson (BF).

Lufttryk under glottis n=80-96 (40-48 ordpar), 1 informant.

Der registreres et noget højere lufttryk under glottis i stødets første fase. I den anden fase ses et fald i trykket under glottis. Det er dog stadig lidt højere end i ord uden stød.



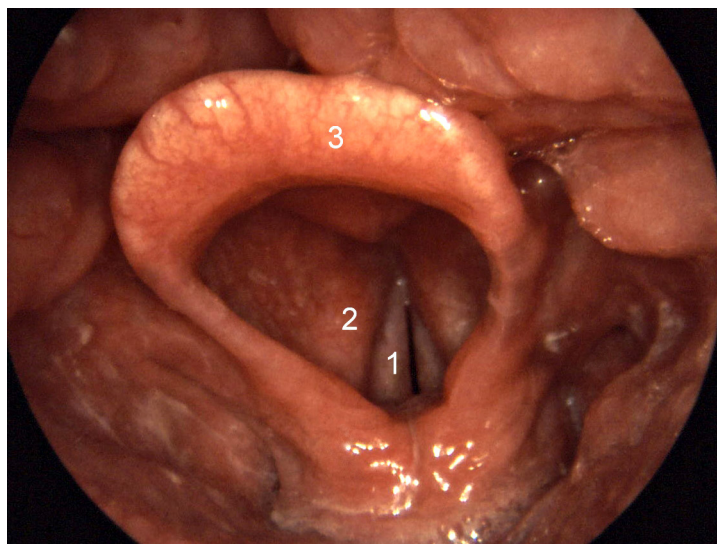
Figur 3.2. Tværnsnit af hoved med angivelse af: tunge (a), velum (drøbel) (b), epiglottis (c), glottis (stimmelæber) (d). Position for fiberskop (1). Positioner for måling af lufttryk: farynx (2), esophagous (spiserøret) (3), under glottis (subglottalt) (4).

Lufttryk i esophagous *n=172 (86 ordpar), 2 informanter.*

I stødets første fase ses et noget højere lufttryk i esophagous (dvs. spiserøret ud for glottis)

Lufttryk i farynx *n= 336 (168 ordpar), 3 informanter.*

I den prævokaliske konsonant ses en svag tendens til at lufttrykket i farynx er højere end i modsvarende ord uden stød. I stødets anden fase ses en tendens til lavere tryk i farynx (over glottis).



Figur 3.3. Foto af stemmekilden, i en vinkel som den ville ses med fiberskop. Stemmелæberne – vocalis (1), falske stemmelæber (2), Strubelåget – epiglottis (3). (Welleschik 2007)

Luftstrøm *n=1132 (566 ordpar), 5 informanter.*

Luftstrømmen gennem munden er målt med en Rothenberg-maske. Der ses en mindsket luftstrøm gennem munden (og dermed gennem glottis) i stavelser med stød, sammenfaldende med stødets anden fase. Luftstrømmen i prævokaliske konsonanter er generelt kraftigere og har en stejlere stigning i stavelser med stød sammenlignet med stavelser uden stød.

Palatogrammer *n= 205, 5 informanter.*

Palatogrammer af vokalen [i] med og uden stød viser en snævrere artikulation (større kontaktflade mellem tunge og gane) i vokaler med stød. Fischer-Jørgensen tolker dette som tegn på større artikulatorisk kraft i stavelser med stød

EMG – Elektromyografisk registrering af muskelaktivitet

Der er foretaget EMG-registrering af muskelaktivitet i vocalis (VOC), interarytenoid (INT), cricothyroid (CT), lateral cricoarytenoid (LCA), posterior cricoarytenoid (PCA) og Thyropharyngeus (TP). Da EMG-resultaterne ikke på nogen tydelig måde bidrager til at forklare stødartikulationen, og da der er tale om vanskeligt stof som det i denne sammenhæng vil kræve uforholdsmæssigt meget plads at formidle på en rimelig måde, tillader jeg mig her at nøjes med at resumere de resultater der vedrører vocalis (VOC) som der vil blive anledning til at vende tilbage til senere i teksten. (Resultater der vedrører PCA og LCA optræder dog i tabel 3.2).

EMG-aktivitet i vocalis (VOC) n=1840-1920 (920-960 ordpar), 9 informanter¹⁵.

VOCs rolle generelt diskuteres på siderne 109-112, før resultaterne der vedrører stødgennemgås. Vocalis er aktiv i forbindelse med forskellige artikulatoriske hændelser:

- Der ses altid aktivitet i vocalis i forbindelse med stemthed (p. 111)
- Der ses ofte, men ikke nødvendigvis, stigende aktivitet i VOC for stigende f_0 (p. 109-110)

"VOC has been found to be active in cases of strong glottal constriction" (p. 111). Det gælder fx:

- Sprængansats (se p. 111 og p. 123)
- Glottal stops (p. 111)
- I forbindelse med synkning (p. 112 nederst)

Ifølge Fischer-Jørgensens fremstilling er det ikke fuldt afklaret hvilken funktion aktiviteten i VOC har i de nævnte tilfælde.

Sekvenser af glottalt lukke og vokal [ʔa ʔa] og vokal med stigende/faldende tone er blevet brugt til at sikre at det rent faktisk er vocalis der registreres. For en enkelt taler ser det ud til at det måske ikke er VOC der er registreret, se pp. 123-126.

Hos syv ud af ni talere registreres en konsistent stærk aktivitet i vocalis-musklen som begynder i vokalens første fase, men som når sit maksimum kort efter det tidspunkt hvor faldet i grundtonen eller uregelmæssighederne i stemmelæbesvingningerne begynder, "thus *preparing* for the second phase" (p. 129, min fremhævnings). .

Selvom det fremgår klart at der er en kobling mellem VOC-aktivitet og stød, fx "[...] peaks in VOC [...] reflect activity for the stød" (p. 119), så er det ikke klart hvad den ekstra aktivitet afstedkommer. Da vocalis som nævnt ovenfor er aktiv i forbindelse med kraftig glottal konstriktion er det oplagt at se på sammenhængen mellem VOC-aktivitet og de øvrige tegn på glottal konstriktion.

Mulige kausale relationer mellem de forskellige registrerede forhold diskuteres i kapitel IV.B.2 pp. 133-142. I forbindelse med diskussionen af årsagerne til det karakteris-

15 984 tokens (492 ordpar) tilskrives én informant; BF.

tiske aftagende intensitetsforløb, som Fischer-Jørgensen betegner som det mest konsekvent forekommende akustiske træk ved stød, fremføres p. 141:

"The decrease in intensity is more difficult to explain. It starts earlier [than the decrease in Fo] (in BF's 1974 recording on the average [43 ms] after the start of the vowel) and gets lower than in vowels without stød after, on the average, [50 ms] (the airflow also starts decreasing early, in BF's 1974 recording, on the average [62 ms] after vowel start). In BRP's inverse filtering curves there is a rather close correlation between decreasing negative spikes, decreasing amplitude of the flow pulses, and decreasing overall intensity. [...] If we assume that BRP, like BF and FJ, does not have much decrease in subglottal pressure, there must be a certain constriction of the larynx at a rather early point. The decrease of oral airflow for the other speakers also points to a larynx constriction. But intensity and airflow start decreasing before the peaks in vocalis and LCA are reached and before the strong constriction of the larynx is seen in the fiberoptic pictures".

Og videre p. 142: "[...] the decrease of intensity starts [64 ms] after vowel onset, i.e. only a little more than [10 ms] later than the rise in VOC activity. However, it often decreases more rapidly later, simultaneously with the Fo decrease".

Timing og forløbet af VOC-aktiviteten passer altså rigtigt ikke sammen med de øvrige indikatorer på glottal konstriktion. Dertil kommer vel også at den konstriktion af de falske stemmelæber, som ses hos en del af talerne, ikke har noget med VOC at gøre, men må tilskrives særskilt muskelaktivitet: "[the] activity in the supraglottal muscles, which must be involved in the adduction of the ventricular folds [...]" (p. 140.)

3.1.3 Opsamling

Generelt er den første fase karakteriseret ved et forholdsvis højt "energiniveau". ("[a] relatively high expenditure of energy" Fischer-Jørgensen 1989a, p. 128). Forholdene i den anden fase står i kontrast til forholdene i den første fase på måder man måske uformelt kunne kalde en fysiologisk og akustisk "opbremsning".

Forud for fase 1

- Luftstrømmen i prævokaliske konsonanter er generelt kraftigere og har en stejlere stigning. Der ses en svag tendens til at lufttrykket i farynx er højere i den prævokaliske konsonant.
- Den EMG-registrerede aktivitet i PCA er højere i prævokaliske ustemte konsonanter.¹⁶

Fase 1

- Altid højere grundtone, ofte højere intensitet. For ca. halvdelen af talerne ses relativt mere energi ved højere formanter (dvs. F4 og F5).
- Lufttrykket under glottis og i esophagous (spiserøret ud for glottis) er noget højere.

Fase 2

- Kraftigt fald i intensitet, men ingen klare tegn på glottalt lukke (med få mulige undtagelser). Uregelmæssige svingninger i 70.8 % af tokens med stød, men med stor variation (8-100 %) mellem talere og mellem forskellige optagelser (sessioner) med samme taler. I 53.8 % af tokens med stød ses et fald i grundtone efterfulgt af eller kombineret med uregelmæssige svingninger. Der ses en tendens til fald i styrke for de lavere deltoner. Invers filtrering viser mindskede "negative spikes" (som afstedkommer mindsket overall intensitet, se videre Fischer-Jørgensen 1989a, p. 75), og den integrerede kurve (af det inverst filtrerede signal) viser faldende glottal luftstrøm og en længere lukkefase.
- De fiberoptiske optagelser af larynx viser konstriktion af stemmelæberne (vocalis) og i mange tilfælde adduktion af de falske stemmelæber (ventricular folds), men begge med stor interindividuel variation. Der ses et fald i trykket under glottis (der dog stadig er lidt højere end i ord uden stød), tendens til lavere tryk i farynx (over glottis) og en lavere luftstrøm gennem munden (og dermed gennem glottis). Det registrerede fald i luftstrøm i den anden fase sammenholdt med et relativt højt tryk under glottis tyder også på en glottal konstriktion.

16 Resultater der baserer sig på et meget lille datamateriale, er her markeret med grå skrift.

Fase 1 og/eller 2

- I distinkt tale er lange vokaler med stød kortere end lange vokaler uden stød, og sonorante konsonanter med stød er længere end modsvarende konsonanter uden stød.
- Palatogrammer af vokalen [i] med og uden stød viser en snævrere artikulation (større kontaktflade) i vokaler med stød. Hos 7 ud af 9 talere registreres en konsistent stærk aktivitet i vocalis-musklen som begynder i vokalens første fase, men som når sit maksimum kort efter det tidspunkt hvor faldet i grundtonen eller uregelmæssighederne i stemmelæbesvingningerne begynder, "thus preparing for the second phase" Der ses et tilsvarende peak i LCA-aktivitet.

Trykgruppen

- I ord med stød ses somme tider et tonalt mønster hvor den tonale top forekommer på den betonedede vokal i modsætning til som ventet (for københavnsk) at toppen forekommer på den første efterfølgende ubetonede stavelse. Fordelingen af forekomsterne af dette omvendte tonale forløb ser ud til at hænge sammen med den betonedede vokals kvalitet og med talerens regionale baggrund.

3.1.4 Diskussion: stød as a phonation type:

Der forekommer ofte (men ikke altid) tydelige ændringer af stemmekvaliteten i forbindelse med stød der gør at den adskiller sig fra stemmekvaliteten i det omgivende sproglige materiale (uden stød). Den eller de måder stemmekilden bruges på i forbindelse med stød må nødvendigvis høre hjemme i det repertoire af mulige almindelige (dvs. ikke patologiske) fonationsmåder som den menneskelige stemme spænder over. Det er derfor naturligt at Fischer-Jørgensen søger at sammenholde sine observationer vedrørende stød med samtidens beskrivelser af forskellige etablerede typer af stemthed. En yderligere anledning er at der er (udenlandske) fonetikere som har beskæftiget sig med stemthedstyper og som derfor har interesseret sig for hvordan stød kunne klassificeres eller beskrives.

I sin diskussion af resultaterne udfolder Fischer-Jørgensen derfor en detaljeret sammenligning af stødet med forskellige relevante fonationstyper. Indledningsvis fastslås det at "[...] of course [...] the first and the second part of a vowel with stød cannot be described as belonging to the same phonation type" (p. 131). De to faser behandles derfor hver for sig.

Fase 1

Forholdene i fase 1 sammenholdes med beskrivelser af tight eller tense voice som det fremstilles af Laver 1980 og Sundberg & Gauffin 1978. Lighederne består i:

- relativt højt subglottalt tryk
- relativt kraftige højere overtoner
- højere intensitet
- mere artikulatorisk kraft¹⁷

Men der er også uoverensstemmelser hvad angår en række andre forhold der nævnes som karakteristiske for tense voice. Det gælder følgende:

- svag f_0 i forhold til F1
- lav luftstrøm (flow) gennem glottis
- glottal konstriktion
- narrow cartilagenous glottis¹⁸
- medial compression

Fischer-Jørgensen fremhæver at disse karakteristika i det store og hele er de samme som nævnes for knirkestemme, og at de ikke er typiske for stødets første fase – måske i særdeleshed set i kontrast til forholdene i stødets anden fase. Fischer-Jørgensen afviser derfor at det kan betegnes som tense/tight voice: "Thus a syllable with stød may be said to start with a relatively strong expenditure of energy, but it is not what is called tense or tight phonation in the literature" (Fischer-Jørgensen 1989a, p. 133)

17 Min oversættelse af "articulatory force". I modsætning til de øvrige forhold på listen er "articulatory force" ikke en målelig størrelse, men en impressionistisk (vel auditivt vurderet) term der relateres til begreberne fortis og lenis.

18 Jeg har ikke kunnet finde en dækkende dansk term, men der skal forstås en indsnævring i den bageste mere bruskholdige del af glottis.

Fase 2

Forholdene i fase 2 sammenlignes med knirk/creak. Også her findes der både ligheds-punkter og forskelle: I de tilfælde hvor der er uregelmæssige stemmelæbesvingninger, ligner det på mange måde knirk (creak) som det beskrives i litteraturen. Det gælder:

- den visuelle fremtræden i spektrogrammer
- lav grundtone (selvom creak ikke nødvendigvis har lav grundtone ifølge Ladefoged)
- svag grundtone (H1) især relativ til niveauet ved F1
- relativt lang lukkefase i det glottale svingningsmønster
- lav luftstrøm (flow) gennem glottis
- periode-til-periode-variation i grundtone (jitter)
- ofte konstriktion i larynx – sammenpressede stemmelæber
- og muligvis adduktion af the falske stemmelæber
- "A steep closure phase in creaky voice was [...] also found by Roach and Hardcastle (1979) and by Esling (1984), and there were indications of this in BRP's flow curves" (p. 132).

Men Fischer-Jørgensen hæfter sig også ved nogle uoverensstemmelser:

- Ifølge den litteratur Fischer-Jørgensen henviser til (Ladefoged, Catford), er det kun den forreste del af stemmelæberne som tager del i svingningerne, mens den bageste del er lukket, eller svinger langsommere (Abercrombie). Det forekommer, men: "It is at any rate not a dominating or a conspicuous feature of the stød" (Fischer-Jørgensen 1989a, p. 132)
- Hovedanken angår dog intensitetsforholdene: "The most consistent acoustic feature of the stød – the decrease of overall intensity – is never mentioned for creak or creaky voice" (p. 132)
- "Catford finds low subglottal pressure in creak" – det er ikke i overensstemmelse med Fischer-Jørgensens resultater for stød.

På den baggrund konkluderer hun at "stød is not just creak". Til det må man tilføje at Fischer-Jørgensen finder at der *ikke* er uregelmæssige svingninger i ca. 29 % af de undersøgte stavelser med stød (se dog bemærkningen ovenfor om bestemmelse af f_0 og om opgørelsen af forekomst af uregelmæssige svingninger). Også af den grund må stød være andet og mere end knirk.

Som nævnt finder Fischer-Jørgensen ingen klare tegn på glottalt lukke, men de fiber-optiske optagelser og registreringerne af luftstrøm og tryk giver klare indikationer på en glottal konstriktion. Stød må derfor høre hjemme i den øvre ende af Ladefogeds kontinuum: "It is evident that the second phase of the stød, the stød proper, belongs in the constricted end of Ladefoged's stricture dimension. Perhaps one might describe an exaggerated stød as a glottal stop, a strong stød as creak, and a milder stød as creaky voice", p. 131. (*Creaky voice* forstås her som et trin mellem modal stemthed og egentlig knirk.)

3.1.5 Diskussion: variability and invariance

I Fischer-Jørgensens staus over de indsamlede data er variabilitet et gennemgående tema. Til trods for det ganske omfattende repertoire af fonetiske metoder der har været bragt i anvendelse, når Fischer-Jørgensen ikke frem til en fonetisk definition af hvad stød er. Fischer-Jørgensen opsummerer meget redeligt de akustiske og fysiologiske forhold der er iagttaget, og fremhæver den store variation der ses, både fra taler til taler og fra token til token:

Fischer-Jørgensen skriver indledningsvis p. 8: "[stød] generally shows up as a decrease in intensity and (often) pitch, in distinct speech ending in irregular vibrations (creaky voice), in very emphatic speech probably sometimes in a complete glottal closure"

Og med fokus på de akustiske resultater, på side p. 146: "The phonetic description of the Danish stød given in this paper, revealed a high degree of variability both in the production and in the acoustic nature of the stød. Not all seem to use their vocalis muscle in the stød, and those who do use it in different ways for other purposes [...] Some have strong adduction of the ventricular folds, some hardly any adduction. Some have a strong pitch decrease, some have rising pitch, and in both cases there is decrease of intensity and there may be irregular vibrations. Nevertheless there is no problem in the perceptual identification [...]"

Om forholdet mellem stød og forekomsten af uregelmæssige svingninger siger hun (p. 52): "Words without stød do not show any irregularity in long stressed vowels or in sonorant consonants (whereas weak syllables with schwa may show some irregularity). Irregular vibrations are thus a characteristic of the stød" Selvom de uregelmæssige svingninger der ses, på mange måder ser ud som knirk, mener Fischer-Jørgensen dog ikke at det er helt det samme: "[...] many features are identical for creak or creaky voice and stød, but it is not quite the same. Stød is not just creaky voice" (p. 133). Det afholder dog ikke forfatteren fra at bruge termen "creak" i forbindelse med stød, se fx citatet på foregående side.

Noget af variationen beskrives som systematisk – fx kobles glottalt lukke i forbindelse med realisering af stød til emfase, og tilstedeværelse af uregelmæssige svingninger kobles til distinkthed – mens noget lader til at afhænge af taleren: Der gives eksempler på talere hvis realisationer mere eller mindre generelt udviser særtræk, som fx fravær af uregelmæssige svingninger, eller stigende tonalt forløb i forbindelse med stød – men i hovedsagen lader variationen til at være tilfældig.

Fischer-Jørgensen må vige uden om at konkludere hvad stød fonetisk set "er". Det fremhæves i stedet at det vigtige er at der opnås perceptuel ekvivalens frem for akustisk eller artikulatorisk invarians: "I agree completely with Lindblom (1982) when he states that what is aimed at is not invariance, but perceptual equivalence. Sufficient, but not necessarily identical acoustic cues are supplemented by top down predictability"

3.2 Grønnum & Basbøll

Grønnum og Basbølls fælles studier har til formål at undersøge om der kan findes empirisk belæg for at sige at dansk har mora. Mora er fonologisk et udtryk for stavelsesvægt (hvormed stavelser kan klassificeres som lette, tunge og overtunge), som ikke traditionelt har været betragtet som relevant for dansk. Basbøll har dog i en række fonologiske analyser af stødet og dets distribution (Basbøll 1988, 1998) taget afsæt i en kobling mellem mora og stød, og har på baggrund af bl.a. Fischer-Jørgensens undersøgelser af stød udtrykt formodninger om at man ville kunne iagttage tegn på mora i den fonetiske substans. Sammenhængen ridses kort op i Grønnum og Basbølls fælles artikler:

"Basbøll (1988, 1998) assumes moraic structure in Danish syllables [...], and thus stød is a property of certain bi-moraic syllables. In phonetic support of his analysis, inter alia, he quotes Fischer-Jørgensen's (1989a) observation that the laryngealization tends to begin about halfway through the vowel, if it is long, or near its offset if it is short, which makes stød-syllables bi-phasal" (Grønnum & Basbøll 2001a). "Vowels with stød are phonologically long; they are proposed also to be bi-moraic, with the second mora carrying the stød" (Grønnum & Basbøll 2002). Se evt. videre Grønnum & Basbøll 2001b, p. 231.

Grønnum & Basbøll 2001a, b, 2002a, b, 2003a, b og 2007 præsenterer tilsammen resultaterne af en række af undersøgelser designet til at afklare om der kan findes fonetisk evidens for mora i dansk. Der er dels tale om akustiske undersøgelser og dels om perceptionsforsøg der har til formål at undersøge sprogbrugernes eventuelle opfattelse af moricitet. (Grønnum & Basbøll 2001a + 2002b er to korte artikler der tilsammen giver et fint overblik over undersøgelserne.)

I Grønnum & Basbøll 2007, p. 199 opsummeres der således:

"Our studies led us to conclude that:

- the exact acoustic properties, the timing, and the segmental domain of the stød are highly variable;
- there is no evidence to indicate that vowels with stød have two perceptually distinct phases.

Basbøll's (1988) proposal, "stød as a signal for the second mora of syllables" is therefore not an acoustic nor an immediate cognitive reality".

Forfatterne konkluderer at stød fonetisk (og perceptuelt) må regnes for en egenskab ved stavelsen.

Perceptionsforsøgene vedrører meget specifikt perceptionen af moricitet. Eftersom de gennemførte undersøgelser ikke giver nogen tegn på at mora skulle være en relevant fonetisk størrelse på dansk vil jeg ikke komme nærmere ind på perceptionsforsøgene.

De akustiske undersøgelser bibringer ny information som har relevans også ud over afklaringen af spørgsmålet om tegn på mora. Undersøgelserne omfatter måling af varigheden af vokaler og konsonanter med og uden stød samt bestemmelse af hvornår laryngaliseringen¹⁹ sætter ind.

3.2.1 Vokalvarighed

Grønnum & Basbøll 2001b finder at der ikke ses en systematisk forskel i varighed mellem vokaler med og uden stød i distinkt tale; (n=350, 175 ordpar). Resultatet afviger til dels fra hvad der tidligere er fundet. Som vi så fandt Fischer-Jørgensen at vokaler med stød er kortere end vokaler uden stød i distinkt tale, mens forskellen ikke opretholdes i mindre distinkt tale (se afsnit 3.1.1). Ifølge Grønnum & Basbøll er grunden til at de ikke når frem til samme resultat som Fischer-Jørgensen 1989a, at det undersøgte materiale ikke er fuldt sammenligneligt: Selvom der i begge undersøgelser er tilstræbt distinkt udtale i det undersøgte materiale, så er Fischer-Jørgensens undersøgelser baseret på ord i citationsform, hvor Grønnum og Basbølls undersøgelse er baseret på oplæste fulde sætninger og "[...] long vowels are relatively longer in words in citation form than in words in natural sentences [...]" (Grønnum & Basbøll 2001b, p. 224). Ifølge forfatterne skyldes det at "Citation form produces focus²⁰, and focus lengthens. However, focal lengthening does not necessarily apply in equal measure to all speech sounds: short vowels cannot be lengthened ad infinitum in a language [...] where vowel length is contrastive. [...]" (Grønnum & Basbøll 2001b, p. 245)

Grønnum og Basbølls fortolkning af forskellen mellem deres og Fischer-Jørgensens resultater må forstås sådan at der kræves både distinkt udtale og fokus før der kan observeres en forskel i varighed mellem vokaler med og uden stød, og at forskellen skyldes at vokaler *uden* stød er længere under fokus i distinkt tale, mens varigheden af vokaler med stød ikke påvirkes af fokus.

19 Laryngalisering betegner her knirk eller i hvert fald uregelmæssige svingninger. Forfatterne ser i praksis ud til at bruge termene creak (knirk), laryngealization og glottalization synonymt.

20 Det formuleres ikke eksplicit i teksten, men "fokus" skal her formodentlig forstås inden for en funktionel-grammatisk tradition. Se fx Dik 1989, p. 277: "Fokus is an abstract function defined in terms of the informational status of the constituent in question within the discourse"

3.2.2 Konsonantvarighed

Grønnum & Basbøll 2001b finder at det afhænger af positionen i ordet og i ytringen om der er en forskel i varigheden af konsonanter med og uden stød:

- ordmedialt ses ingen forskel (n=210, 5 informanter)
- ordfinalt finder de at konsonanter med stød er 10-30 ms længere end konsonanter uden stød (n=228, 5 informanter)
- ytringsfinalt finder de at konsonanter med stød er 20-50 ms kortere end konsonanter uden stød (n=107, 5 informanter)

Den største forskel findes altså i ytringsfinal position, og forfatterne fremhæver at "utterance-final consonants with stød stand out in their shortness" (Grønnum & Basbøll 2001b, p. 241). Forfatterne konkluderer at der ikke er basis for at sige at konsonanter med stød *generelt* er længere end tilsvarende konsonanter uden stød.

Også dette resultat adskiller sig fra hvad der tidligere er fundet: Som vi så fandt Fischer-Jørgensen 1989a (i overensstemmelse med Riber Petersen 1973) at konsonanter med stød var længere end modsvarende konsonanter uden stød. Grønnum og Basbøll rejser i første omgang tvivl om hvorvidt der er tale om en reel forskel i resultaterne: Efter deres vurdering må der sættes spørgsmålstegn ved Riber Petersens resultater da de ser tegn på at der har været vanskeligheder med segmenteringen (dvs. afgrænsningen) af konsonanterne på grund af de uregelmæssige svingninger. De henleder opmærksomheden på at Fischer-Jørgensen skriver at netop segmenteringen af konsonanter med stød var vanskelig, og at målingerne af konsonantvarigheder i Fischer-Jørgensen 1989a derfor også er mindre sikker.

Som nævnt adskiller Grønnum og Basbølls materiale sig fra det der ligger til grund for tidligere undersøgelser, ved at bestå af oplæsning af hele reelle sætninger. Som supplement har de dog også optaget nogle lister af ord (n = 14, 7 ordpar). Disse er ikke medregnet i de resultater jeg refererede indledningsvist. I dette begrænsede materiale som er mere direkte sammenligneligt med det materiale Fischer-Jørgensen har analyseret, viser konsonanter med stød sig at være længere end modsvarende konsonanter uden stød – i overensstemmelse med Fischer-Jørgensen og Riber Petersens resultater. Det tyder på at konsonanter med stød *kan være længere* i citationsform (eller under fokus).

3.2.3 Timing af stødfasen (laryngalisering)

Ifølge Grønnum og Basbøll er stød "a kind of creaky voice", og stødet, eller stødfasen, kan derfor afgrænses tidsmæssigt ved at se på hvornår der forekommer laryngalisering – dvs. uregelmæssigheder i stemmelæbesvingningerne.

Starttidspunktet for laryngaliseringen (dvs. overgangen til uregelmæssige stemmelæbesvingninger) kan ikke bestemmes i ca. 32 % af det undersøgte materiale (n = ca. 557, 5 talere). Der skelnes ikke mellem tilfælde hvor tidspunktet for overgangen fra

modal stemthed til laryngaliseret ikke kan fastsættes, og tilfælde uden synlig laryngalisering. Der er ikke nogen opgørelser over i hvor stor en andel af ord med stød der forekommer synlig laryngalisering, men det er klart at andelen må være på mindst 68 %.

I den del af materialet hvor laryngalisering er til stede, og hvor timingen af den kan bestemmes, ses følgende:

- Laryngaliseringen starter i gennemsnit 66-68 ms efter vokalstart, med nogen variation i gennemsnitsværdierne fra taler til taler (fra 36 ms til 73 ms).
- Set over materialet som helhed varierer starttidspunktet for laryngaliseringen betydeligt: fra 10 til 130 ms regnet fra vokalstart.
- Der ses ingen nævneværdig forskel i timingen af laryngaliseringen imellem vokaler og konsonanter med stød.
- Der ses ingen (eller kun ringe) korrelation mellem timingen af laryngaliseringen og vokalvarighed.
- Forfatterne hæfter sig dog ved at der er en forskel i gennemsnitsværdierne for de snævre af de undersøgte vokaler [i:][y:] sammenlignet med den mere åbne vokal [e:] på henholdsvis 62 ms og 78 ms, der nøje modsvares af en tilsvarende forskel i (intrinsic) varighed, som er på henholdsvis 115 ms og 135 ms ($p < 0.01$ for begge). Mulige forklaringer på dette forhold diskuteres.

Disse to sidste observationer kunne synes at være i modstrid med hinanden. Eftersom den beregnede korrelation er lav, må man i hvert fald formode at forskellen i timing af laryngaliseringen mellem de snævre og den ikke snævre vokal er overlejret med en betydelig variation.

Da formålet med at se på timingen af laryngaliseringen er at se om der kan findes akustiske tegn der markerer grænsen mellem de foreslåede mora, er der fokuseret på hvornår de uregelmæssige svingninger begynder. Der fremlægges ikke systematiske opgørelser over hvor lange stræk der er laryngaliserede, eller hvornår laryngaliseringen ophører relativt til det stødbærende segments afslutning. Det fremføres dog flere gange at laryngaliseringen kan forstærke ud over varigheden af det stødbærende segment, fx Grønnum & Basbøll 2001b, p. 234 "[...] stød in a long vowel may extend into a succeeding sonorant consonant. Stød in a sonorant consonant may already begin during the last part of the preceding short vowel and may continue into a succeeding voiced sound as well. Nor does the stød phase, the creaky voice, have a very definite and fixed duration". Tilsvarende udsagn ses i Grønnum & Basbøll 2007, p. 199 "[...] the variable timing of the occurrence of creaky voice [...] and the variable total duration which often makes it continue well into the following syllable"

3.2.4 Varighedsforhold

Det forhold at vokaler uden stød i distinkt tale er længere under fokus, mens vokaler med stød ikke er længere under fokus, kalder på en forklaring. Ifølge Grønnum og Basbøll må forklaringen på at vokaler med stød ser ud til at have samme varighed uanset om de står i fokus eller ej, søges i det forhold at stød-fasen ikke kan forlænges.

At stød-fasen ikke kan forlænges, bygger de dels på deres intuition, dels på en observation de har fra Dorte Bleses; forfatterne fremfører: "Dorte Bleses [pers. commun] has often observed, in child-directed speech, that when mothers – for whatever reason – lengthen vowels that would ordinarily have stød, the stød is absent. This corroborates with our own intuition that it is not possible to draw out the stød phase at will" Den forskel der findes i Fischer-Jørgensen 1989a "[...] would be due to a focally induced lengthening of the long stødless vowel, rather than to any shortening effect of the stød" I umiddelbar forlængelse følger dog denne overvejelse: "This line of reasoning unfortunately leaves unexplained, and even contradicts, the longer (intervocalic) stød consonant in citation form [...]" (Grønnum & Basbøll 2001b, p. 245).

Uanset at varighedsforholdene for vokaler og konsonanter med stød ikke tegner et samstemmende billede, fastholder Grønnum og Basbøll ideen om at det er stødet der begrænser muligheden for at forlænge vokaler med stød.

3.2.5 Bemærkninger til varighedsforhold

Grønnum og Basbølls resultater viser at laryngaliseringens begyndelse er relativt løst knyttet til det stødbærende segments begyndelse, og de mere uformelle observationer vedrørende laryngaliseringens sluttidspunkt tyder på at den er endnu mindre knyttet til det stødbærende segments afslutning sådan at laryngaliseringen gerne "flyder ud over" segmentgrænsen. Det kan derfor virke ejendommeligt at stød-fasen skulle indvirke på varigheden af en stødbærende vokal sådan at stødet (eller stød-fasen) skulle hindre at vokalen forlænges.

Da varighedsforholdene for konsonanter med og uden stød ikke følger det samme mønster som varighedsforholdene for vokaler med og uden stød, forekommer det under alle omstændigheder ikke rimeligt at indfortolke en kausal relation mellem stød og varighed: Man må nøjes med at konstatere at vokaler med stød ikke forlænges under fokus, mens konsonanter forlænges under samme omstændigheder.

I tilknytning til disse overvejelser må man også bemærke sig at Dorte Bleses førmtalte observation ikke kan bruges til at vise at det skulle være umuligt at forlænge (vokaler med) stød. Observationen illustrerer blot at der i disse tilfælde sker noget andet. Det kan måske forstås sådan at forlængelsen kan overtrumfe behovet for stød under visse omstændigheder.

3.2.6 Grønnum og Basbølls ballistiske model

Grønnum og Basbølls observationer giver dem anledning til at foreslå en ny model for stød:

"We have faith in our intuition and in Dorthé Bleses' observation that the stød phase cannot be lengthened at will. Many other speech gestures can be controlled in time, why not stød? Perhaps the gesture which controls its production is ballistic in nature. This is actually not an unreasonable assumption, given Fischer-Jørgensen's [1987]²¹ investigation: 5 of her 7 speakers had a very clear vocalis muscle activity in syllables with stød (the remaining 2 subjects *may* employ a different strategy, unless the electrode was simply not in the exactly right position in the muscle). This is how she describes it [Fischer-Jørgensen, 1987, p. 163, VOC is short for 'vocalis'].

The extra activity in VOC in words with stød begins in most cases 2-4 cs after the start of the vowel [...]. It has its peak 10-13 cs after vowel start [...], i.e. shortly after the point where the stød phase (with decrease in frequency or irregular vibrations) starts [...]. The activity decreases quickly again and crosses the curve for the corresponding word without stød approximately at the end of the vowel for long vowels with stød and at the end of the consonant in the cases of short vowel with stød in the following consonant [...]. The activity is thus generally of rather brief duration (10-20 cs in the average curves and often somewhat shorter in individual examples), and the peak may be very sharp.

The sharp rise and fall – with no intermediate plateau – in the EMG traces is characteristic of every single illustration in the article, and is very reminiscent of a low-pass-filtered response to a transient source" (Grønnum & Basbøll 2001b, p. 248)

Modellen udfoldes yderligere i Grønnum & Basbøll 2007, pp. 199-200:

"Fischer-Jørgensen (1987, 1989a, b) presents an account of various phonation types, particularly creaky voice. She concludes that many features are identical for creaky voice and stød, but that stød is not just creaky voice. We agree entirely. On the basis of our acoustic and perceptual observations, here is how we propose to characterize the stød phonetically:

- the laryngeal activity is a ballistic gesture which, minimally, makes for a slightly compressed voice quality, at one end of a continuum, and, maximally, creates a distinctly creaky voice at the other. Under emphasis it may become a complete glottal closure;
- it is a property of the syllable rhyme as a whole;
- it is aligned with the onset of the rhyme;
- it is variable with respect to strength and to temporal extension.

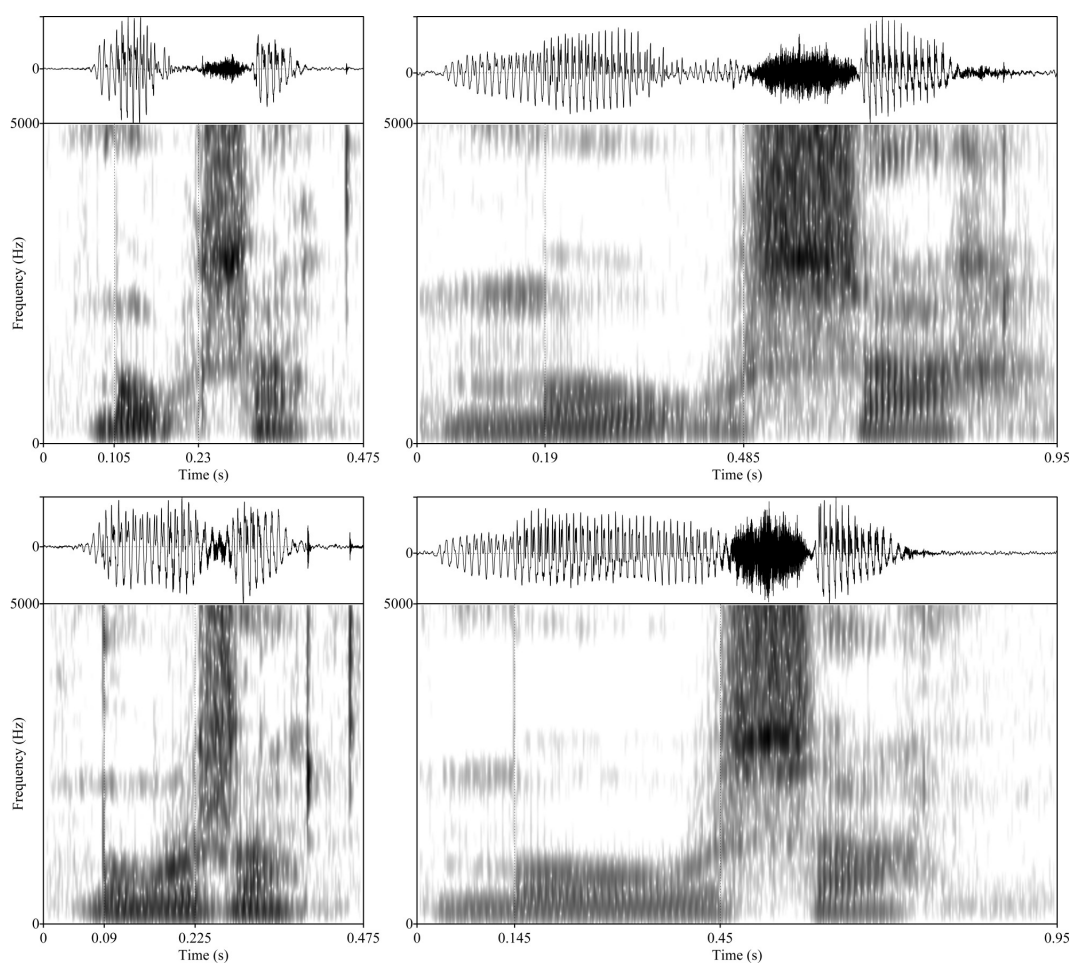
21 Fischer-Jørgensen 1987 er den tidligere udgave af Fischer-Jørgensen 1989a.

The proposed ballistic gesture is to be understood as the low-pass filtered muscular response to a transient neural command. The neural command is presumably timed to coincide with the onset of the rhyme. The impulse may be stronger or weaker, resulting in more or less irregular vocal fold vibration of shorter or longer duration, but once the command is executed, the speaker can no longer control the way the vocal folds respond to the excitation, just as one can no longer control the trajectory of a cannon ball once the cannon is fired. This proposal is consistent with the fact that speakers cannot freely choose to increase the duration of the stød the way one may choose to lengthen creaky voice at the end of an utterance. It is also consistent with the way we have seen the stød to behave acoustically: the more or less explicit non-modal vocal fold vibration, the variable timing of the occurrence of creaky voice in the waveform and the spectrogram, and the variable total duration which often makes it continue well into the following syllable. Furthermore, our proposal is consistent with EMG-data: the higher vocalis muscle activity in stød relative to modal voice increases and decreases gradually (Fischer-Jørgensen 1987, 1989a, b). It would be curious indeed if the actual mechanical change in vocalfold vibration mode were not also gradual"

3.2.7 Bemærkninger til Grønnum & Basbølls model

Præsentationen af den ballistiske model indledes med formuleringen: "We have faith in our intuition and in Dorte Bleses' observation that the stød phase cannot be lengthened at will" (Grønnum & Basbøll 2001b, p. 248).

Det er værd at bemærke, at udsagnet om at stød ikke kan forlænges, er en antagelse. Som tidligere fremført (3.2.5) kan Dorte Bleses observation ikke bruges til at vise at det skulle være umuligt at forlænge (vokaler med) stød. Jeg vedstår også gerne at min intuition ikke stemmer overens med Grønnum og Basbølls: Jeg tror det *er* muligt at variere varigheden af stød – i hvert fald inden for visse rammer. Som eksempel og som en uformel afprøvelse af min formodning vedføjes optagelse af et minimalpar med og uden stød (moser, mo'ser) sagt i hurtigt og langsomt tempo. (Lydfilen findes som appendix A.VII). Oscillogram og spektrogram af de to ordpar ses i figur 3.4.



Figur 3.4. Eksempler på minimalpar med og uden stød (øverst: mo'ser, nederst: moser) sagt hurtigt (tv) og langsomt (th). Varighed af vokal med stød er hhv. 125 ms og 295 ms. Varighed af vokal uden stød er hhv. 135 ms og 305 ms.

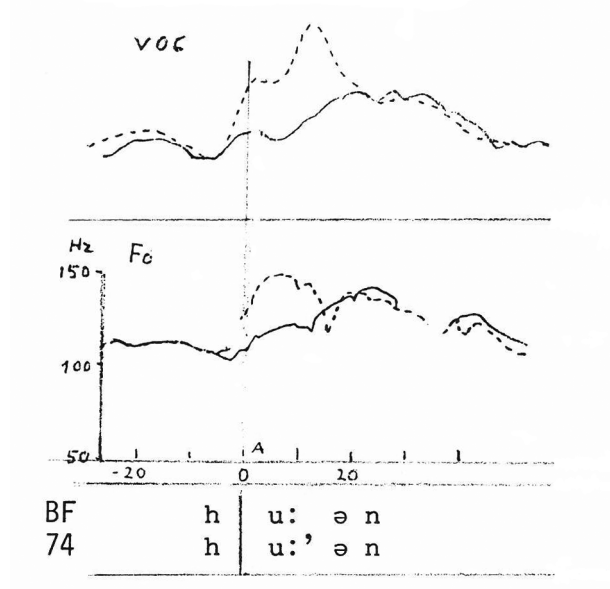
Optagelserne illustrerer at der er muligheder for at realisere en tydelig forskel i varighed af både vokalen med stød og i det man må betragte som de akustiske udslag af stødet; en passage med mindre regelmæssige svingninger og med mindsket amplitude.

Min påstand er ikke at de to udtaler repræsenterer "almindeligt forekommende tale", men det er dog mulige realisationer af ordene, og en præsentation af minimalparrene for kolleger bestyrker mig i den opfattelse at distinktionen mellem de to ord oprettholdes både i den hurtige og den langsomme fremføring. Jeg medgiver at det muligvis ikke er almindeligt forekommende at forlænge vokaler med stød²², fx i forbindelse med tøven, men jeg kan ikke dele den opfattelse at det skulle være artikulatorisk eller fysiologisk umuligt.

²² Det er så vidt vides ikke undersøgt systematisk endnu.

Som det fremgår finder Grønnum og Basbøll støtte for deres model i de EMG-data som præsenteres i Fischer-Jørgensen 1989a: "The sharp rise and fall – with no intermediate plateau – in the EMG traces is characteristic of every single illustration in the article [...]". Kigger man nærmere på de pulicerede EMG-kurver i Fischer-Jørgensen 1989a, vil man dog se at det ikke forholder sig helt så enkelt som fremstillingen i Grønnum & Basbøll 2001b giver indtryk af.

Tager man som eksempel den første figur med EMG-data for VOC i Fischer-Jørgensen 1989a (figur 37 i Fischer-Jørgensen 1989a; udvalgte dele gengivet nedenfor som figur 3.5), ser man i EMG-forløbet for VOC først en stejl stigning i ordet med stød (stiplet linje) umiddelbart op til vokalens start (markeret med en lodret linje), dernæst et plateau der ser ud til at omfatte omtrent 1/3 af vokalens varighed. Herefter følger endnu en stejl stigning (der ikke er helt så kraftig som stigningen op til vokalens start), før niveauet igen falder relativt hurtigt til et niveau som er lidt lavere end plateauet i begyndelsen af vokalens.

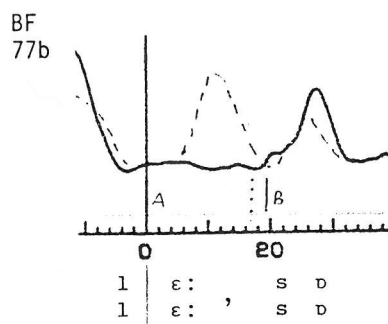


Figur 3.5. EMG-aktivitet i vocalis og grundtone (f_0) i hhv. [hu:'en] (stiplet linje) og [hu:en] (fuldt optrukket linje). Den lodrette linje markerer vokalens begyndelse.

Ved ordenes begyndelse er VOC-niveauet det samme i ordet uden stød som i ordet med stød. I ordet uden stød (fuldt optrukket linje) stiger VOC betydeligt mindre op til vokalens begyndelse, og der hvor den anden stigning sætter ind i ordet med stød, ses i ordet uden stød en langsommere og svagere stigning der når sit maksimum og flader ud ved omtrent samme niveau og på samme tidspunkt som VOC-niveauet i ordet med stød falder og flader ud efter peaket.

Overordnet set beskriver EMG-kurverne relativt komplekse forløb både i ordet med og i ordet uden stød. EMG-forløbet i [hu:'en] udviser desuden et tydeligt plateau i vokalsens første del.

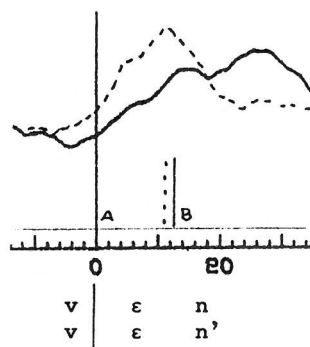
Der ses også markante peaks i stavelser uden stød. Fx viser figur 3.6 [læ:'sp]/[lɛ:'sp] (figur 1b, appendix IX, Fischer-Jørgensen 1989a) et peak i anden stavelse for begge ord som er af næsten samme størrelsesorden som det peak der ses i forbindelse med stødet i første stavelse.



Figur 3.6. EMG-aktivitet i vocalis i hhv. [læ:'sp] (stiplet linje) og [lɛ:'sp] (fuldt optrukket linje). Den lodrette linje ved A markerer vokalsens begyndelse. De lodrette linjer ved B markerer afslutningen på vokalen i hhv. ordet med stød (stiplet linje) og ordet uden stød.

I ordet med stød (stiplet linje) ses først et stigende-faldende forløb der når sit maksimum omtrent midt i vokalen i første stavelse. Der ses et tilsvarende stigende-faldende forløb, men med et mindre udsving i den efterfølgende ubetonede stavelse. I ordet uden stød (fuldt optrukket linje) ses et meget jævnt niveau igennem den første stavelse og et stigende-faldende forløb i anden stavelse som er kraftigere end det der ses i ordet uden stød i samme stavelse. Det er dog ikke helt så kraftigt som det der ses i første stavelse af ordet med stød.

Den registrerede EMG-aktivitet i vocalis i ord med stød er ikke lige kortvarig eller transient i alle tilfælde. I figur 3.7 der viser ordparret [ven']/[ven] (figur 1a, appendix IX, Fischer-Jørgensen 1989a) ses i ordet med stød (stiplet linje) et noget langsommere stigende-faldende forløb sammenlignet med de foregående eksempler. Forløbet når sit maksimum omtrent ved overgangen mellem vokalen og den sonorante konsonant. I ordet uden stød (fuldt optrukket linje) starter forløbet ved et lidt lavere niveau i [v]'et og stiger langsommere til et niveau som er lidt lavere end i ordet med stød; det maksimale niveau nås i sidste del af den sonorante konsonant.

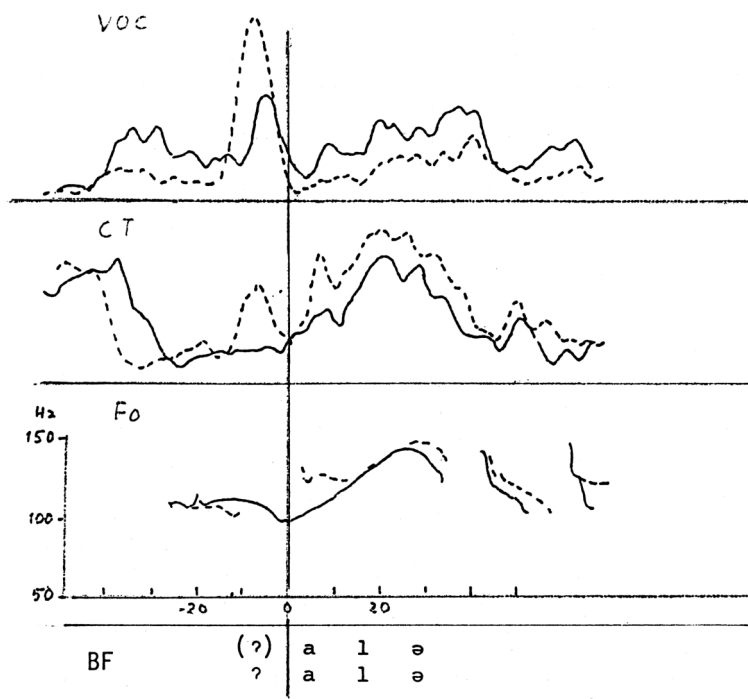


Figur 3.7. EMG-aktivitet i vocalis i hhv. [ven'] (stiplet linje) og [ven] (fuldt optrukket linje). Den lodrette linje ved A markerer vokalens begyndelse. De lodrette linjer ved B markerer afslutningen på vokalen i hhv. ordet med stød (stiplet linje) og ordet uden stød.

De kraftigste EMG-signaler i vocalis er ikke registreret i forbindelse med stød. Fischer-Jørgensen fremfører p. 173: "VOC and LCA were also found to be active for hard attack [dvs. sprængansats] for BF, and in this case the activity is often still stronger than for the stød" (med reference til figur 45 i Fischer-Jørgensen 1989a, gengivet her som figur 3.8).

I figur 3.6 ses ved begyndelsen af eller umiddelbart forud for [l]'et i ordet uden stød højere EMG-aktivitet end der ses i vokalen i ordet med stød. Da den foregående stavelse ud fra beskrivelsen af materialet er ubetonet, kan denne kraftige EMG-aktivitet ikke have noget med stød at gøre. Det er uvist hvad der er årsag til det høje niveau i dette tilfælde, men der ville udmærket kunne forekomme sprængansats på dette sted.

Forløbene i figur 3.6 og 3.7 er desuden eksempler på at EMG-aktiviteten i VOC i ordene uden stød kan nå omtrent samme niveau som peaket i ordene med stød. I disse tilfælde ses peaket ikke nær en stavelsesgrænse som ville kunne give anledning til en sprængansats, så der er ikke nogen umiddelbar forklaring på det høje niveau.



Figur 3.8. EMG-aktivitet i vocalis og cricothyroidæus samt grundtone (f_0) i ordet [(?)alə] udtalt med og uden sprængansats (hard attack) ved vokalens begyndelse. Stiplet linje markerer udtale med sprængansats, fuldt optrukket linje markerer udtale uden sprængansats. Den lodrette linje markerer vokalens begyndelse.

Fischer-Jørgensen giver selv udtryk for at der må være en forbindelse mellem VOC-aktivitet og stød, fx: "It is apparent from these examples that activity in the VOC is closely connected with the stød" (Fischer-Jørgensen 1989a, p. 171), men det er også klart at forekomsten af kraftig og/eller transient aktivitet i VOC ikke er begrænset til stavelser med stød.

Da der ses løbende aktivitet i VOC med vekslende intensitet med både langsomme og forholdsvis abrupte forløb, og da der ikke kun forekommer kraftig og/eller transient aktivitet i VOC i forbindelse med stød men også andre steder, så er der for mig at se ikke grund til at tolke den aktivitet der ses i forbindelse med stød, som tegn på en særlig styringsmekanisme sådan som Grønnum og Basbøll har foreslået.

3.3 Knirk – fra uønsket epifænomen til definerende træk

Et af de karakteristiske træk ved stød er de ofte forekommende uregelmæssige svingninger – som oftest benævnt knirk. Fischer-Jørgensen og Grønnum og Basbølls beskrivelser af stød adskiller sig fra hinanden med hensyn til hvilken rolle de uregelmæssige stemmelæbesvingninger gives i beskrivelsen. Kaster man blikket lidt længere bagud i tid, tegner der sig konturerne af en interessant udviklingshistorie; læser man fx Svend Smith 1944, kan man finde følgende syn på forekomsten af uregelmæssige svingninger i forbindelse med stød:

"Man taler fra stemmehygiejnisk Side [dvs. inden for talepædagogikken] om, at Stødet kan udføres mere eller mindre godt, idet man har henledt Opmærksomheden paa det knirkede Stød som noget, der ikke hører hjemme i en Bedstenorm" I tilfælde hvor der *ikke* forekommer tab af periodicitet (dvs. uregelmæssige stemmelæbesvingninger) "[...] kan man tale om normativt Stød. Er der derimod Tab af Periodicitet [...] maa der dertil svare en knirkende Lyd (rettere Ikke-Tone), og vi taler om et ikke-normativt Stød [...]" (Smith 1944, p. 98).

Ifølge Smith bør stød altså i den bedste af alle verdener realiseres uden uregelmæssige svingninger²³. I Smiths egne oscillografiske undersøgelser af stød forekommer der dog uregelmæssige svingninger i ca. halvdelen af ordene med stød (n= ca. 106 ord med stød, 5 informanter). Fordelingen er dog meget uens sådan at tre informanter stort set konsekvent udviser uregelmæssige svingninger i forbindelse med stød, mens én informant kun har uregelmæssige svingninger i et ud af 10 tokens med stød. I Svend Smiths egen udtale, der udgør ca. halvdelen af de undersøgte tokens, ses uregelmæssige svingninger i 15 ud af 57 ord med stød, svarende til ca. 26 %.

Fischer-Jørgensen kommenterer direkte på Smiths normative tilgang og siger at selv om man ud fra en talepædagogisk synsvinkel måtte mene at det ikke er en anbefalelsesværdig måde at frembringe stød på, så forekommer det ikke desto mindre så ofte at det må betragtes som helt almindeligt: "The most conspicuous acoustic characteristic of words with stød is the presence of irregular vibrations in the latter part of a long vowel or in a sonorant consonant following a short vowel . i.e. in what Smith calls the "second phase" of the stød, generally considered as the real stød phase. Smith considers this as a "non-normative" stød because, as a speech therapist, he does not find this way of pronouncing the stød recommendable. But as a matter of fact this is quite a normal pronunciation of the stød in distinct speech . In the present investigation irregularities were found in 70.8 % of all cases" (Fischer-Jørgensen 1989a, p. 50)

Fischer-Jørgensen bemærker desuden at hvor der ofte forekommer uregelmæssige svingninger i forbindelse med stød, så ses der i det undersøgte materiale ikke uregelmæssige svingninger i betonedede vokaler uden stød. "Words without stød do not show any irregularity in long stressed vowels or in sonorant consonants (whereas weak

23 Synspunktet lader til at være almindeligt for perioden, og Smith er her på linje med talepædagoger som fx V. Forchhammer. Se fx Forchhammer 1954, pp. 180-181.

syllables with schwa may show some irregularity). Irregular vibrations are thus a characteristic of the stød" (Fischer-Jørgensen 1989a, p. 52).

Andelen af stød der udviser uregelmæssige svingninger, varierer en del i Fischer-Jørgensens undersøgelser: "[...] the largest differences are found between different readings by the same speaker. [...] It seems mainly to be a question of the distinctness level chosen by the speaker in the session in question"

Selvom Fischer-Jørgensen kalder de uregelmæssige svingninger for: "The most conspicuous acoustic characteristic of words with stød" så er det andre træk hun giver forrang i sin karakteristik af stød: "[stød] generally shows up as a decrease in intensity and (often) pitch, in distinct speech ending in irregular vibrations (creaky voice) [...]" (Fischer-Jørgensen 1989a, p. 8).

I Grønnum og Basbølls fremstilling rykker de uregelmæssige svingninger frem i første række: "Stød is a kind of creaky voice, a laryngealization which characterizes certain syllable rhymes²⁴ under certain conditions" (Grønnum & Basbøll 2001b). Laryngalisering eller (en slags) knirk, i praksis identificeret ved tilstedeværelse af synlige tegn på uregelmæssige svingninger (i det akustiske signal), indtager her en position som det definerende træk ved stød²⁵.

Det er denne arbejdsdefinition²⁶ der er lagt til grund for Grønnum og Basbølls fælles undersøgelser. Som nævnt giver deres undersøgelser anledning til at foreslå en alternativ forklaring. Motivationen er først og fremmest at kunne forklare forhold vedrørende stødets udtrækning i tid, men samtidig giver den foreslåede model mulighed for at håndtere det ellers paradoksale forhold at auditivt tydelige stød der ikke udviser tegn på uregelmæssige svingninger (og som Grønnum og Basbøll selv behandler som stød), principielt falder uden for den anvendte (arbejds)definition.

I den model der præsenteres i Grønnum & Basbøll 2007, skifter de uregelmæssige svingninger status fra at være det definerende træk til at være en blandt flere mulige konsekvenser af en ballistisk laryngal gestus: "[...] the laryngeal activity [...] minimally, makes for a slightly compressed voice quality, at one end of a continuum, and, maximally, creates a distinctly creaky voice at the other"

Det ses således at de famøse uregelmæssige svingninger over tid (og til dels på tværs af faggrænser) er blevet tilskrevet ret forskelligartet status. Hvor de hos Smith og i den talepædagogiske tradition ses som et uønsket epifænomen, figurerer de hos Grønnum og Basbøll, i hvert fald en overgang, som det definerende træk: stød er (en slags) knirk.

24 I fonologisk analyse kan stavelser ses som opdelt i onsets og rim. Rimdelen kan underopdeles i nukleus og koda; "stol" får da onsets "st", rim "ol", som består af nukleus "o" og koda "l".

25 Variationer over denne definition findes foruden i Grønnum og Basbølls fælles artikler i Grønnums grundbøger i fonetik (Grønnum 2001, 2007) og i en oversigt over dansk fonetik i JIPA (Grønnum 1998: Illustrations of the IPA : Danish)

26 Min formulering.

4 Ny viden om knirk og en hypotese om stød

4.1 Stød fonetisk set – status quo

Hvor der på det fonologiske niveau hersker en ganske klar opfattelse af hvornår der kan tales om stød, så er der ikke på tilsvarende måde vished eller enighed om hvad stød er rent fonetisk set – hverken akustisk eller artikulatorisk.

Der er udbredt enighed om at stød er et glottalt fænomen, og Fischer-Jørgensens resultater peger ret tydeligt på at der forekommer en glottal konstriktion i forbindelse med stød. Der er samtidig enighed om at stød ikke generelt kan sidestilles med glottalt lukke, men at kraftige stød, fx under emfase, *kan* realiseres som – eller indebære – et glottalt lukke. Fischer-Jørgensen nævner en række karakteristiske akustiske og fysiologiske træk, men vil ikke lægge sig fast på en fonetisk karakteristik af stødet.

Grønnum og Basbøll har brugt definitionen "Stød er en slags knirk", se fx afsnit 3.2.3. I forlængelse af deres fælles studier forkaster Grønnum og Basbøll den definition og formulerer i stedet et forslag til en ny måde at karakterisere stød fonetisk. Den ballistiske model de har foreslået, er dog ikke underkastet en empirisk afprøvning.

Status quo er således at der ikke eksisterer en etableret og veldokumenteret definition på eller karakteristik af hvad stød er fonetisk set.

4.2 Knirk eller ej?

Uregelmæssige svingninger fremhæves ofte som et prominent akustisk træk ved prototypiske stød selvom det ikke tillægges samme vægt i alle fremstillinger.

De uregelmæssige svingninger omtales ofte som knirk, men Fisher-Jørgensen konstaterer akustiske og fysiologiske forhold i forbindelse med de uregelmæssige svingninger der forekommer i forbindelse med stød, som ikke harmonerer med samtidens beskrivelser af knirk, og når på den baggrund frem til at der ikke kan være tale om genuin knirk. Grønnum og Basbøll på deres side taler om stød som "en slags knirk".

De forskelle Fisher-Jørgensen påpeger, imellem den type uregelmæssige svingninger som forekommer i forbindelse med stød sammenlignet med samtidens beskrivelser af knirk, må betyde at uregelmæssige svingninger ikke altid er af samme art. Netop det forhold kaster en artikel af Janet Slifka ("Some Physiological Correlates to Regular and Irregular Phonation at the End of an Utterance") fra 2006 nyt lys over: Emnet for Slifka 2006 er at karakterisere de uregelmæssige svingninger der meget ofte ses sidst i ytringer (med udgangspunkt i Amerikansk Engelsk). I forbindelse med almindelig såkaldt modal brug af stemmen åbner og lukker stemmelæberne for luften i en jævn vekslen sådan at der er åbent og lukket omtrent lige stor del af tiden. I visse modeller for stemmekildens funktion kaldes dette forhold for "Open Quotient" (OQ forklares nærmere i afsnit 5.2). Overkomprimeret eller presset stemmekvalitet og knirk er karakteriseret ved en forholdsvis længere lukket fase og en kortere åben fase, dvs.

lavere OQ, og ved en mere abrupt åbning og sammenklapning af stemmelæberne. Knirk er derudover karakteriseret ved uregelmæssigheder i timingen af åbne- og lukkefaserne (eller ved meget lav grundtone, lidt afhængig af hvilken beskrivelse man tager udgangspunkt i). Slifkas pointe er at de uregelmæssige svingninger der ses ytringsfinalt, er karakteriseret ved det modsatte forhold mellem længden af den åbne og den lukkede fase: altså lang åben fase og kort lukket fase, dvs. høj OQ. Undersøgelsen viser at stemmelæberne kan vibrere uregelmæssigt på mere end en måde der dårligt kan betegnes som knirk. Da knirk bl.a. er defineret ved lav OQ, så kan de uregelmæssige svingninger der typisk ses ytringsfinalt og som altså viser sig at have en høj OQ, ikke betegnes som knirk. Slifka demonstrerer med tydelighed at der er mere end én måde at skabe uregelmæssige stemmelæbesvingninger på, og man kan derfor ikke sætte lighedstegn mellem uregelmæssige stemmelæbesvingninger og knirk.

Det er ikke alle der betragter uregelmæssige svingninger som et definerende træk ved knirk. Det gælder fx Catford 1964 der bare taler om meget lav grundtone. Når der forekommer uregelmæssige svingninger, lader det dog til at have været almindelig praksis at betragte det som knirk. På baggrund af Slifkas undersøgelse må man derfor konstatere at man ikke kan være sikker på at tidligere undersøgelser af "knirk" alle er undersøgelser af samme eller sammenlignelig substans.

Der er derfor grund til at genoverveje om de uregelmæssige svingninger der ofte forekommer i forbindelse med stød, kan karakteriseres som knirk eller ej.

Hvis knirk er karakteriseret ved høj kompression (og dermed lav OQ), uregelmæssige svingninger og glottal indsnævring²⁷, så peger det i retning af at antage at der rent faktisk ér tale om knirk i forbindelse med stød når der ses uregelmæssige svingninger, idet Fischer-Jørgensen finder klare tegn på glottal konstriktion i forbindelse med stød og samtidig tegn på en mere komprimeret brug af stemmen. Disse tegn er dels at de spektrale analyser viser mere energi ved højere formanter, og dels Fischer-Jørgensens pilotforsøg med invers filtrering der tydede på en lavere OQ.

Fischer-Jørgensen har dog følgende invending: "[...] the most consistent acoustic feature of the stød – the decrease of overall intensity – is never mentioned for creak or creaky voice" (p. 132). Selvom der almindeligvis ses en sammenhæng mellem lydstyrke og kompression, der viser sig ved at kompressionen normalt er højere jo højere der tales, så er det ikke sådan at en bestemt stemtheds måde automatisk resulterer i et bestemt lydtryk. Det er derfor sandsynligvis sådan at knirket stemthed og modal stemthed inden for visse rammer kan produceres med samme styrke.

Ser man på Ladefogeds hierarki af stemtheds måder (se tabel 4.1 i afsnit 4.5) er det klart at den maksimale styrke af den lyd som stemmekilden kan generere, må være faldende mod begge yderender af hierarkiet, altså henholdsvis glottalt lukke og helt åben glottis. Den stemtheds måde der giver mulighed for den højeste intensitet, er formentlig noget komprimeret, men ikke knirket; altså omtrent midtvejs i hierarkiet.

27 Jævnfør Lindqvist 1969, 1972a og 1972b, som Fischer-Jørgensen også refererer til.

Når stød indebærer en bevægelse i retning af "the constricted end of Ladefogeds continuum" (Fischer-Jørgensen 1989a, p. 131), så forekommer det ikke urimeligt at tænke sig at det samtidig kunne medføre et dyk i intensiteten.

Fischer-Jørgensen peger også på at registreringen af relativt højt subglottalt tryk i forbindelse med stød ikke harmonerer med målinger rapporteret af Catford, der tværtimod finder lavt subglottalt tryk i forbindelse med knirk. Fischer-Jørgensen tager det som endnu en faktor der indikerer at der ikke kan være tale om knirk i forbindelse med stød. Men forskellen kunne også opstå hvis det forholdt sig sådan at det var Catfords data der ikke var baseret på egentlig knirk. Sammenholder man Catfords og Slifkas resultater kunne det godt se ud til at være tilfældet: I Slifkas undersøgelse af uregelmæssige svingninger observeres lavt glottalt flow sammen med høj OQ, dvs. lille glottal modstand. Deraf kan man udlede at det glottale tryk må have været relativt lavt. Det forekommer derfor meget sandsynligt at Catfords data kunne være baseret på uregelmæssige svingninger af samme type som dem Slifka har undersøgt, og dermed ikke "knirk" i traditionel forstand.

4.3 Forklaringsbehov I: Hvad er stød fonetisk set?

Der er ikke tvivl om at uregelmæssige svingninger er et prominent træk ved stød i sin prototypiske form. Selvom Slifkas undersøgelse sandsynliggør at de uregelmæssige svingninger der forekommer i forbindelse med stød, med rette kan benævnes som knirk, så er jeg helt enig med Fischer-Jørgensen og med Grønnum og Basbøll i at stød ikke *kun* er knirk:

Det er ikke al knirk der har med stød at gøre, og ikke alle stød udviser knirk. Tilstedeværelsen af knirk er hverken tilstrækkeligt eller nødvendigt som kriterium for at der er tale om stød.

4.4 Forklaringsbehov II:

Hvordan kan "synlige" og "usynlige" stød forekomme lige tydelige?

Der er ikke i sig selv noget overraskende i at et fonem²⁸ kan realiseres og genkendes uden at dets mest karakteristiske eller definerende træk er til stede. Det overraskende ligger i at det skulle forekomme så ofte i meget distinkt tale.

Det gælder for stort set alle fonetiske størrelser at de kan realiseres mere eller mindre fuldbyrdet, og at mindre distinkte realisationer i almindelig brug uproblematisk kan opfattes til trods for at de kun indebærer antydninger af de primære træk. Som eksempel kan nævnes lukkelyde i intervokalisk position der kan reduceres til approximant: "ikke" → [eʊə]. Men selvom den fonologiske valeur kan opretholdes med en reduceret fonetisk realisationen, så vil man (i hvert fald til dels med øvelse) kunne høre forskel på de mere eller mindre reducerede realisationer, og man vil i mange tilfælde opleve en forskel i hvor tydelig den fonetiske størrelse er at høre.

Hvis uregelmæssige svingninger er et prominent træk ved stød i sin prototypiske form, så ville man forvente at finde dette træk nogenlunde konsekvent i den mest distinkte talestil.

Man ville også forvente at tilfælde med mindre eller slet ingen uregelmæssigheder ville blive opfattet som mindre tydelige eller reducerede instantieringer af stød – hvilket ikke synes at være tilfældet. Der er ikke eksperimentelt belæg for at stød uden knirk generelt er vanskeligere at høre eller at de opfattes som reducerede. Tværtimod kan eksempler som fx Grønnum og Basbøll har fremdraget (fx Grønnum & Basbøll 2007, p. 196, eller Grønnum 2005, p. 215) tages som tegn på netop det modsatte: at i hvert fald visse stød uden uregelmæssige svingninger eller andre akustiske tegn der forbindes med stød, høres som tydelige, fuldgylde, ikke-reducerede stød.

Det er det forhold at tokens der høres som utvetydige, tydelige og distinkte eksempler på stød, ikke nødvendigvis udviser hvad der betragtes som de primære akustiske karakteristika, der er påfaldende.

28 Stød betragtes ganske vist som en prosodi og ikke som et fonem, men stød kan være betydningsadskillende og dermed have fonematisk funktion.

4.5 Ny forklaringsmodel for stød

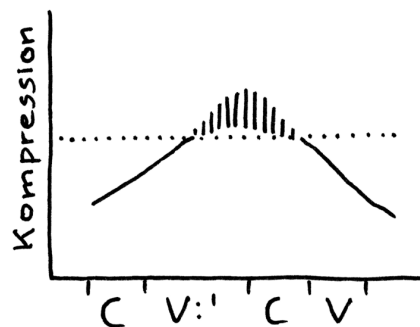
I det følgende vil jeg foreslå en måde at indpasse den labile tilstedeværelse af knirk i en model for stød uden at labiliteten i tilstedeværelsen knyttes til mere eller mindre distinkt udtale. Jeg tager udgangspunkt i en traditionel rangordning af stemmekvaliteter ud fra den glottale åbningsgrad som den præsenteres hos Ladefoged (1971), se tabel 4.1

Åbningsgrad	Fonetisk term	Kompression
lukket	glottal stop	
.	creak	
.	creaky voice	
.	pressed voice	↑ Overkomprimeret Hyperfunktionel ~ lav OQ
.	modal voice	• Ideal / optimal
.	breathy voice	↓ Underkomprimeret Hypofunktionel ~ høj OQ
.	whisper	
åben	breathing	

Tabel 4.1. Stemmekvaliteter ordnet efter glottal stricture (åbningsgrad). Med udgangspunkt i Ladefoged 1971.

Hypotesen er at stød udtrykkes gennem en relativt kortvarig forandring af stemmekvaliteten i retning af en presset evt. knirket stemmekvalitet og efterfølgende tilbage til mindre presset stemmekvalitet²⁹. Modellen er illustreret i figur 4.1. Stød ansues dermed som en dynamisk stemmekvalitetsgestus. Et velformet stød tænkes at indebære et passende stort udsving i stemmekvaliteten over et passende (kort) tidsrum. Om der forekommer knirk i forbindelse med stød eller ej, tænkes at afhænge af hvor højt oppe på skalaen stødforløbet ligger.

²⁹ Det er muligt at betragte hypotesen som en formalisering af den beskrivelse af stød som man finder hos V. Forchhammer. Se Blegvad (red.) 1954, p. 180-181.



Figur 4.1. Illustrering af det hypotetiserede kompressionsforløb i ord med stød; i dette tilfælde lang vokal med stød. Den stiplede linje illustrerer en tænkt grænse mellem regelmæssige og uregelmæssige stemmelæbesvingninger. Den fuldt optrukne op- og nedadgående kurve illustrerer hhv. stigende og faldende kompression, og passagen med lodrette streger angiver uregelmæssighed i stemmelæbernes svingninger. Bredden af forløbet i den uregelmæssige del illustrerer desuden at kompressionen nok ikke kan bestemmes med samme sikkerhed – måske endda slet ikke – når signalet ikke er periodisk.

Konceptuelt kan der drages paralleller til tonale forløb som fx det der udspiller sig omkring betonedede stavelser i dansk eller ordtoner i fx svensk. Det der er afgørende for at viderebringe de tonale forløb, er netop heller ikke at nå op eller ned til en bestemt tonehøjde, men at der udspilles et bestemt tonalt forløb over et passende grundtoneområde og at dette times ganske nøje i forhold til de berørte stavelser. Hvor trykgruppemøsteret og ordtonerne udfolder sig i det tonale domæne, tænkes stødet at udfolde sig i et klangmæssigt domæne.

4.6 Forbehold og implicitte antagelser

Ladefogeds kontinuum *er* en forsimpning af forholdene der ikke rummer alle muligheder. Fx er luftfyldthed og kompression delvist uafhængige; en stemme kan godt være komprimeret og luftfyldt (se fx Reinholt 2008), men ofte går de hånd i hånd sådan at en luftfyldt stemme også er underkomprimeret. Tilsvarende viser Slifkas undersøgelse at uregelmæssige stemmelæbesvingninger ikke nødvendigvis fordrer høj kompression, sådan som det ellers er den almindeligt udbredte opfattelse.

Der er også andre svagheder ved det opstillede kontinuum. Det er fx ikke rigtig klart hvor diplofoni hører hjemme i systemet af stemmekvaliteter. Jeg antager tentativt at diplofoni hører hjemme mellem egentlig knirk og presset stemmekvalitet.

Selv om man er opmærksom på at hierarkiet ikke kan favne alle forekommende stemmekvaliteter, så går grundtanken i hierarkiet igen i beskrivelser af stemmekilden og dens funktionsmåder. Det gælder også modeller for stemmekildens funktion, som præsenteret i fx Stevens 2000, eller Fant et al. 1985, og mange af de akustiske mål der relaterer sig til stemmekvalitet, som fx $H1:H2$ ³⁰ (se videre kapitel 5).

30 Traditionelt skrives $H1-H2$. Jeg har valgt formen $H1:H2$ for at betone at der er tale om en

Selv om man måske kunne få det indtryk af Ladefogeds kontinuum at der skulle være tale om et endimensionelt forhold, så skal man nok snarere opfatte det som en mere eller mindre etableret ligvistisk relevant rute igennem et flerdimensionelt akustisk og artikulatorisk landskab. Gordon & Ladefoged lægger ikke skjul på at der er tale om en simplificeret fremstilling: "Although this is somewhat of an oversimplification, there nevertheless appears to be a linguistic continuum that can be characterized using these terms as an ordered set" (Gordon & Ladefoged 2001).

4.7 Forklaringspotentiale – muligheder og begrænsninger

Den foreslåede model har først og fremmest potentiale til at forklare den labile tilstedeværelse af uregelmæssige svingninger og hvorfor stød uden knirk ikke nødvendigvis opfattes som reducerede: Det der afgør graden af distinkthed, antages at være størrelsesordenen af bevægelsen op og ned af skalaen og ikke hvor langt op ad kompressionsskalaen man når. Et mindre distinkt stød tænkes altså at være et stød der realiseres med en forholdsvis lille ændring i kompressionsgrad, mens et tydeligt stød realiseres med en kraftigere forandring. Dermed vil der naturligvis også være større sandsynlighed for at der forekommer en passage med uregelmæssige svingninger i stød der realiseres med et stort udsving i kompression.

Det er ikke al den variation der forekommer i realisationen af stød, som den foreslåede model kan forklare. Det gælder fx de tonale forhold, om det nu er dyk i grundtone, eller det omvendte trykgruppemønster som Fischer-Jørgensen beskriver. Modellen gør heller ikke noget forsøg på at forklare de forhold vedrørende varighed som Grønnum og Basbøll jævnfør afsnit 3.2 fremhæver.

styrkerelation mellem første og anden deltone.

4.8 Eksplicitering af forskelle mellem de to forklaringsmodeller

Grønnum og Basbølls forslag vedrører først og fremmest stødets artikulation, men de fremhæver også at stød kan være tydelige at høre selvom de ikke er tydelige at se i akustiske analyser. Den model de foreslår, rummer mulighed for at forklare hvordan der kan forekomme forskellige realisationer, nemlig på baggrund af en kraftigere eller svagere underliggende neural kommando. Modellen lægger ikke op til en forklaring af hvordan det kan være at realisationer der akustisk er meget forskellige, kan opleves som auditivt lige tydelige.

Den model jeg har foreslået, vedrører både produktion og perception: Hvis stød kan anskues som en dynamisk stemmekvalitetsgestus – nemlig et relativt kortvarig ud-sving i stemmekvaliteten i retning af en presset evt. knirket stemmekvalitet og tilbage igen – så må det også antages at det er denne dynamiske ændring i stemmeklangen vi perciperer som signal for stød. Heraf følger også hvad det er for akustiske fællestræk der skal ledes efter for at binde de forskellige realisationsformer sammen, nemlig akustiske korrelater for hvor komprimeret stemmen er.

De to modeller bygger på forskellige ideer om hvordan styringsmekanismerne fungerer: Grønnum og Basbøll foreslår at det er en transient neural kommando der afstedkommer en ballistisk laryngal gestus. Jeg foreslår at der er tale om en særlig måde at bruge en alment tilgængelig resource – stemmekvalitet – til at skabe et særligt dynamisk forløb. Bort set fra netop dette aspekt er de to forslag gensidigt uafhængige, og de udelukker ikke hinanden.

4.9 Kort om hvordan hypotesen kan testes

Den foreslåede model kan testes relativt enkelt med akustiske metoder der kan følge eventuelle klangforandringer, dvs. ændringerne i stemmekvalitet der måtte forekomme i forbindelse med stød. Ideelt set bør stemmekvaliteten følges hele vejen igennem stødet, men hvis det ikke viser sig at være muligt fordi passager med mere eller mindre uregelmæssige svingninger formodentlig vil medføre nogle måletekniske begrænsninger, så må stemmekvaliteten i det mindste følges op til og lige efter eventuelle passager med uregelmæssige svingninger. Emnet for de to følgende kapitler vil være en afsøgning af potentielt brugbare akustiske metoder.

5 Måder at kvantificere (aspekter ved) stemmekvalitet

Hypotesen om stød som en stemmekvalitetsgestus og den tænkte kobling med Ladefogeds kontinuum af stemmekvaliteter er formuleret på et abstraktionsniveau som gør at den ikke direkte kan underkastes en empirisk afprøvning. Det er nødvendigt at eksplicite hvordan Ladefogeds kontinuum relateres til modeller for stemmekildens funktion som lyd giver og dermed modeller for det akustiske signal som genereres.

Der er derfor behov for en kort indføring til en lidt mere detaljeret beskrivelse af hvordan stemmekilden fungerer. Specifikt er der behov for indblik i hvordan forløbet af stemmelæbernes åbne- og lukke-faser influerer på den spektrale sammensætning af den lyd som kilden genererer. Eftersom lyden fra stemmekilden ikke kan optages direkte, er der også behov for at introducere kilde-filter-modellen.

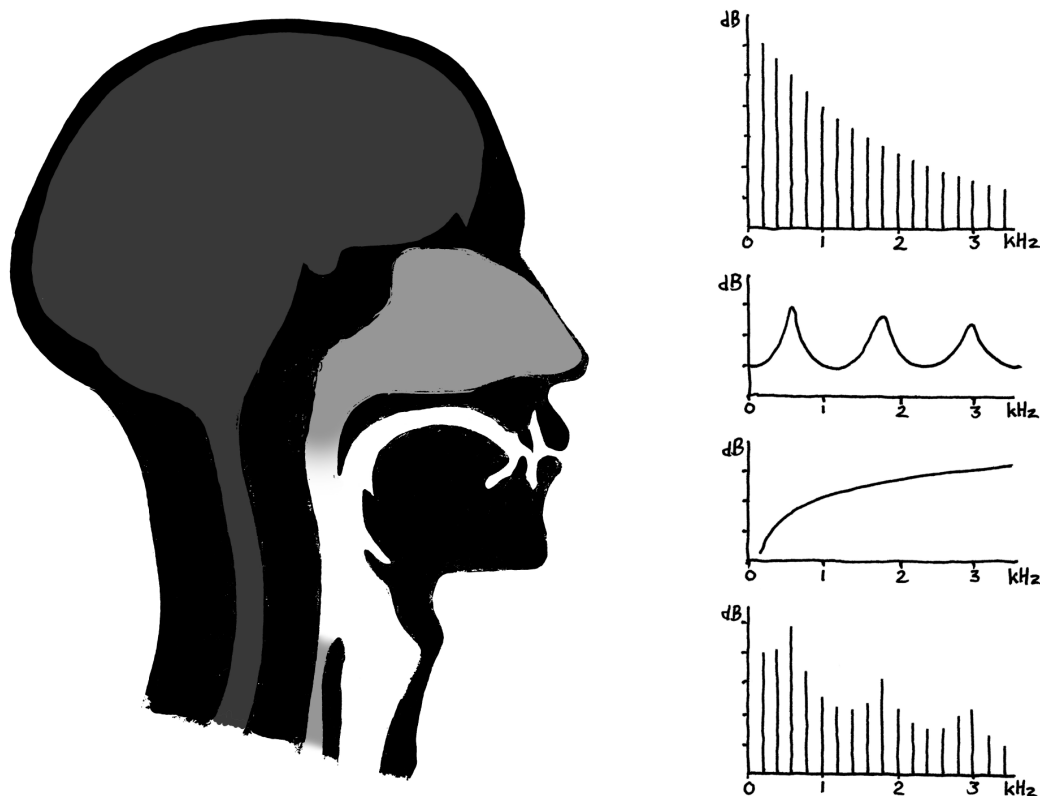
I det følgende vil jeg derfor kort oprids kilde-filter-modellen, sammenhængen mellem måden stemmelæberne svinger på, og de konsekvenser det har for kildespektrets sammensætning. Jeg vil kort relatere beskrivelsen til modeller for stemmekildens funktion (nærmere bestemt LF-modellen; Fant et al. 1985) og beskrive hvordan OQ, som er en af parametrene i LF-modellen, kobles til Ladefogeds kontinuum

Målet er at undersøge en eventuel sammenhæng mellem stød og stemmekvalitet, dvs. klangforandringer i lydsignalet som kan henføres til stemmekilden. Vanskeligheden består i at det talesignal vi kan høre – og optage – er et produkt af taleorganerne som et samlet hele; lyden fra stemmekilden kan ikke høres direkte. Jeg vil derfor forklare noget om de grundlæggende muligheder og begrænsninger der er ved at foretage analyser på baggrund af det akustiske signal der kan optages. Ved at benytte kilde-filter-modellen kan man i princippet regne baglæns fra det signal der er optaget til de bagvedliggende delkomponenter. Det er imidlertid forbundet med både teoretiske og praktiske vanskeligheder der gør at jeg i stedet vælger at afsøge mulighederne for at se på de spektrale forhold i outputspektret. Ved at sammenligne matchede par af ord med og uden stød er det muligt at inferere sig frem til forskelle i spektret der må tilskrives forandringer i kildespektret i forbindelse med stød.

Endelig vil jeg pege på de vanskeligheder man må forvente at støde på i forbindelse med uregelmæssige eller diplofoniske svingninger, og beskrive hvordan man kan forsøge at imødegå dem.

5.1 Kilde-filter-modellen

For analytiske formål kan det være frugtbart at anskue den menneskelige produktion af talelyde som et system bestående af en eller flere lydgivere og et akustisk filter. Talekanalen fungerer med sine (fleksible) hulrum som resonansrum, dvs. som et variabelt akustisk filter der former signalet fra lyd giveren.



Figur 5.1. Illustration af kilde-filter-modellen. De fire grafer illustrerer (fra oven); kildens spektrum, talekanalens filterfunktion, udstrålingsfunktionen (dvs. den filterfunktion som de akustiske forhold omkring hovedet bidrager med), og spektret for den lyd der er det samlede resultat af de tre delkomponenters bidrag.

Denne måde at anskue det akustiske system bag produktion af talelyde kendes som kilde-filter-modellen (Fant 1960, resumeret i Stevens 2000, p. 127-130):

$$\text{kilde} \cdot \text{filter}^{31} = \text{output}$$

For stemte sproglyde er det stemmelæberne der fungerer som den primære lyd giver. For ikke stemte sproglyde kunne det være friktionsstøj dannet ved turbulens i forbindelse med en snæver passage i talekanalen der udgør det lyd givende system.

31 Man betragter ofte talekanalens og udstrålingsfunktionens samlede filterfunktion under ét.

5.2 Stemmekilden – funktionsmåder og modelering

Som skitseret i afsnit 2.4 varierer stemmekildens akustiske spektrum med luftstrøms-funktionen; dvs. de fluktuationer i luftstrømmen som genereres af den kontinuerte vekslen mellem åben og lukket glottis.

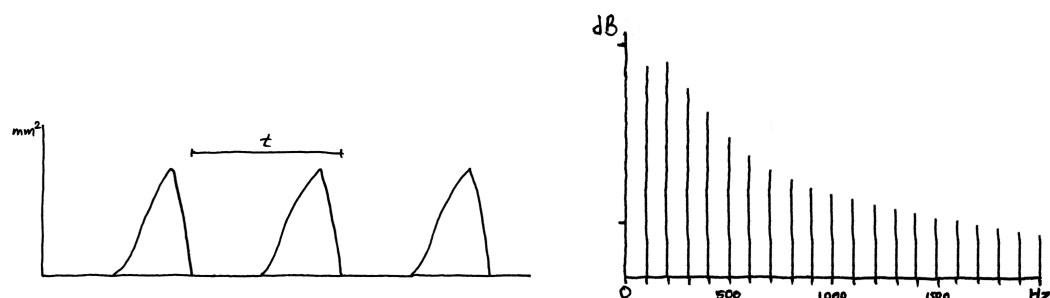
Der kan opstilles modeller for hvordan stemmekilden fungerer som kan redegøre for stemmekildens væsentligste muligheder som lyd giver, se fx Stevens 2000, pp. 97-100 der giver en oversigt over de væsentligste modeller. To væsentlige parametre i modellerne er Open Quotient (OQ) der betegner forholdet mellem den tid i hver periode hvor glottis er åben og den tid glottis er lukket og T2 der udtrykker hvor hurtig den sidste del af lukningen af glottis foregår.

Ved at variere OQ kan man beskrive et spektrum af stemtheds måder der stemmer fint overens med centrale dele af Ladefogeds kontinuum. Med modellerne kan det eftervises at en komprimeret stemmekvalitet med kortere åben-fase (OQ) og en mere abrupt lukning vil have et spektrum med mere energi over grundtonen, og omvendt at en underkomprimeret stemme med en længere åben-fase (OQ) og eventuel glottal lækage (dvs. luftfyldt stemme) vil have et spektrum med relativt mindre energi over grundtonen. (Se videre Stevens 2000, p. 55-100, særligt p. 86, 90 og 99)

Ved modelering vil man se at ændringer i OQ påvirker den nedre del af spektret og dermed giver sig udslag i ændrede H1:H2-forhold (se Fant 1995, p. 127; formel 9 og figur 7), mens T2 har betydning for ved hvilken frekvens spektret overgår fra en 6 dB slope til en 12 dB slope – dvs. groft taget kilde signalets spektrale tilt. Se Stevens 2000, p. 98. Der viser sig at være en så stærk kovariation imellem parametrene³² at man til analytiske formål ofte fokuserer på OQ.

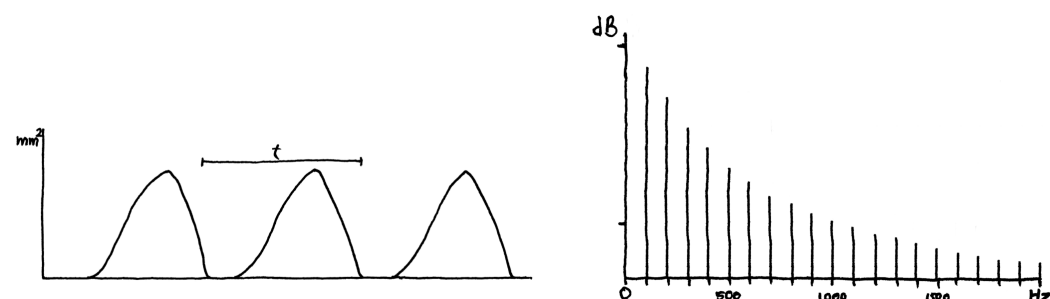
³² På baggrund af en statistisk og funktionel analyse af kovariationen foreslår Fant 1995 en enkelt "shape parameter" benævnt R_d .

Ved typisk modal stemthed vil forholdet mellem stemmelæbernes åbne og lukkede fase være ca. 1:1. Lyden fra stemmekilden vil have sit højeste energiindhold omkring grundtonen (H1) og den første overtone (H2), som kan være omtrent lige kraftige, og derover en overtonerække med gradvist aftagende styrke: Den spektrale profil vil være faldende opefter med ca. 12 dB pr. oktav.³³ (Se figur 5.2).



Figur 5.2. Skitsering af glottal åbning sum funktion af tid (tv.) og det resulterende spektrum (th.) for modal stemthed.

Ved *hypofunktionel* stemthed (underkomprimeret stemme) varer stemmelæbernes åbne fase længere end ved modal stemthed. Den lukkede del er tilsvarende kortere, og selve lukningen foregår mindre abrupt. Stemmekildens spektrale profil vil aftage med mere end 12 dB pr. oktav, dvs. falde mere stejlt. Der vil altså være mindre energi ved højere deltoner (se figur 5.3). Sammenlignet med modal stemthed bliver klangen mere afrundet fordi grundtonen dominerer mere.

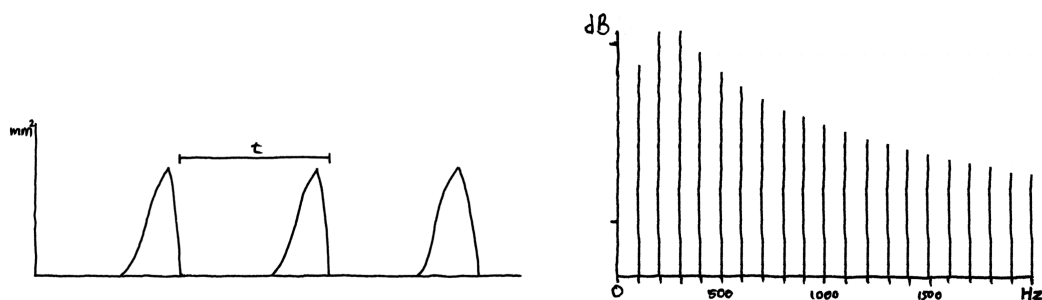


Figur 5.3. Skitsering af glottal åbning sum funktion af tid (tv.) og det resulterende spektrum (th.) for underkomprimeret (hypofunktionel) stemthed.

³³ De præcise værdier for forholdet mellem den åbne og den lukkede del af forløbet, og kildespektrets nøjagtige udseende varierer noget imellem forskellige fremstillinger (og forskellige modelleringer) og vil heller ikke være ens for kvinde- og mandestemmer (se fx Stevens 2000, p. 66 og 72), men beskriver dog i alle tilfælde de samme overordnede tendenser: Mere eller mindre faldende tendens for spektret og særlige forhold for styrkeforholdet mellem de nederste deltoner. De glottale forløb i figur 5.2-5.4 er skitseret på baggrund af Stevens 2000, kapitel 2. Kildespektrene er baseret på Johnson 1997, p. 129.

Ved *hyperfunktionel* stemthed (overkomprimeret stemme) varer stemmelæbernes åbne fase kortere tid end ved modal stemthed. Den lukkede fase er tilsvarende længere, og selve lukningen foregår mere abrupt.

Stemmekildens spektrale profil vil aftage med mindre end 12 dB pr. oktav. Ved hyperfunktionelle (komprimerede) stemmer kan grundtonen (H1) have lavere styrke end de første overtoner. I figur 5.4 gælder det H2-H4. Sammenlignet med modal stemthed bliver klangen skarpere fordi overtonerne har relativt mere energi i forhold til grundtonen.



Figur 5.4. Skitsering af glottal åbning som funktion af tid (tv.) og det resulterende spektrum (th.) for overkomprimeret (hyperfunktionel) stemthed.

Hvis stød som foreslået realiseres som en ændring af kompressionen, må man forvente at det giver sig udslag som løbende forandringer i kildespektret i stavelser med stød.

5.3 Akustisk analyse: beskrivelse af kilde og filterfunktionen

Når man har en optagelse af et passende talesignal (dvs. fx vokallyde), så kan man med udgangspunkt i kilde-filter-modellen i princippet beregne den ene af de to faktorer hvis man har viden om den anden. Kender man talekanalens filterfunktion, så kan man ved at filtrere optagelsen med den inverse filterfunktion rekonstruere lyden fra stemmekilden. Eller omvendt, hvis man kender kildens spektrale karakteristik, så kan man ved at sammenholde den med det registrerede spektrum bestemme filterfunktionen.

Der er dog en alvorlig hage ved sagen: Problemet i forhold til at basere en analyse på det akustiske signal er at kilde-filter modellen giver os én ligning med to ubekendte. Outputtet kan optages som et lydsignal, men hverken kildefunktionen eller filterfunktionen kan registreres direkte. Det betyder at man står overfor et principielt uløseligt problem: Der vil være uendeligt mange kombinationer af tænkte kilder og filtre som vil kunne resultere i det samme output.

I praksis betyder det at den eneste farbare vej når man vil analysere et givet output, er at gøre antagelser om den ene komponent og på den baggrund beregne den anden.

5.3.1 Akustisk analyse: estimering af filterfunktionen

Det går generelt godt når man vil estimere filterfunktionen. Det skyldes hovedsagelig at kildens spektrale karakteristika er enklere og mindre variabel end filterfunktionen. Desuden kan centrale aspekter ved filterfunktionen – formanternes frekvensmæssige placering – ses pr. øjemål i et spektrogram. Det giver en indikation af hvor talekanalens resonansfrekvenser er centreret, selvom man ikke fuldtud kan bestemme filterfunktionens karakteristika på baggrund af et spektrogram.

En hyppigt anvendt måde at estimere filterfunktionen på er ved såkaldt Linear Prediction Coefficient-analyse; LPC-modellen bygger på en forsimplet model for resonanssystemet som kun indeholder poler (resonansfrekvenser hvor der sker forstærkning), men ingen nuller (frekvenser hvor der sker dæmpning). Metoden beregner automatisk det all pole-filter som bedst muligt redegør for forskellen mellem et tænkt kildesignal og det registrerede samlede output (dvs. lydoptagelsen). Man går almindeligvis ud fra at kildefunktionens spektrum har et fald på 12 dB pr. oktav. Formantfrekvenserne og deres båndbredder beregnes så de så godt som muligt redegør for de resterende spektrale bakker og dale. Det har den naturlige konsekvens at hvis den konkrete kildefunktion afviger fra den antagne model, så vil estimatet af talekanalens filterfunktion naturligvis afvige fra den reelle filterfunktion.

5.3.2 Akustisk analyse: estimering af kildefunktionen – invers filtrering

At lave den komplementære analyse – altså at beregne kildesignalet ud fra en optagelse på baggrund af en formodning om filterfunktionen – er en vanskeligere opgave, der ikke kan gøres med tilsvarende præcision.

Når man beregner kildesignalet ud fra det samlede output, så sker det ved at applicere det inverse filter på outputtet – altså et modfilter som ophæver talekanalens filterfunktion. Det er teoretisk muligt at gøre så længe der i det oprindelige filter (talekanalen) ikke er nogen frekvensområder hvor dæmpningen er uendelig stor hvilket typisk forårsages af destruktiv interferens. Er der frekvenser som er blevet dæmpet fuldstændigt, kan det oprindelige signal ved disse frekvenser naturligvis ikke gendannes. Fordi talekanalens filterfunktion varierer mere og på en mere kompleks måde end kildefunktionen gør, er det forbundet med større vanskeligheder at estimere kildefunktionen.

Ved at bruge viden om talekanalens konfiguration ved forskellige vokalkvaliteter og dermed de resonansfrekvenser som kan beregnes på baggrund af akustiske modeller, og ved at have en god formodning om hvordan kildesignalet må se ud (på baggrund fx af fysiologiske modeller af stemmekilden og auditive bedømmelser af den konkrete optagelse), kan man, især i forbindelse med statiske (holdte) vokaler, modellere talekanalens filterfunktion tilstrækkelig nøjagtigt til at det har ført til frugtbare undersøgelser af kildens spektrale karakteristika. Se fx oversigt i Fant 1993, og jævnfør Fischer-Jørgensens pilotforsøg som blev refereret i afsnit 3.1.1, s. 29.

I praksis foregår analysen ved en interaktiv process hvor man på baggrund af den opfattede vokalkvalitet indstiller det inverse filter, og derefter finjusterer det indtil resultatet er et plausibelt kildesignal. Det vurderer man ud fra den spektrale energifordeling, udseendet af det hypotetiserede lydsignal ved glottis og den deraf beregnede luftstrøm gennem glottis.

Selv med disse indsnævringer i mulighedsrummet er der dog stadig vide muligheder for at finde plausible kombinationer af kilde og filterkarakteristika som resulterer i det samme output. Invers filtrering efter den fremgangsmåde er med andre ord ikke nogen eksakt metode³⁴. Samtidig er det en meget tidskrævende process – selv når det handler om at undersøge "statiske" vokaler hvor grundtone og formanter holdes konstant³⁵. Løbende tale giver yderligere udfordringer fordi filterfunktionen vil være under konstant forandring.

Man kunne få den tanke at man kunne bruge LPC-analyse til at finde den filterfunktion der skal til for at lave invers filtrering. Men sådan en fremgangsmåde ville være cirkulær: Eftersom LPC-analysen går ud fra en bestemt model af kildefunktionen, så betyder det at en invers filtrering lavet på baggrund af en LPC-analyse ville ende med at blive en beskrivelse af den model af kildefunktionen som LPC-analysen er baseret på.

34 Fant 1995: "Conventional LF-parameter extraction from inverse filtering is far from an exact procedure and is subject to variations in individual strategies and, thus, runs the risk of being more an art than a science.

35 Et eksempel på en "statisk" vokal kunne være et langstrakt "øøh".

5.4 Holistisk akustisk analyse : muligheder og problemer

Selvom det ville være ønskeligt at analysere direkte på det akustiske signal fra stemmekilden, så fungerer kilde og filter set fra en sproglig vinkel som en enhed; Fant 1995, p. 120 "It is the combined gesture, rather than the source function alone, which has a communicative function". Det forekommer derfor rimeligt at når stød kan høres og genkendes, uanset de forskellige omgivelser det findes i, så må det også have akustiske egenskaber som kan måles i det lydlige output der umiddelbart kan høres og optages. Men med forventningen om at signaleringen af stød hovedsagelig må henregnes til kildefunktionen, så er det klart at det ikke kan være absolutte akustiske mål man skal kigge efter i outputspektret. Problemet ved at forsøge at kvantificere klangforskelle mellem vokaler (eller stavelser) med og uden stød som kan relateres til stemmekilden på baggrund af akustiske målinger af outputspektret, er naturligvis at outputspektret også afhænger af eventuelle samtidige forskelle i vokalkvalitet, tonehøjde og generel stemmekvalitet.

Man kan imødegå problemet ved at sammenligne analyser af matchede minimale ordpar som så vidt muligt kun adskiller sig ved tilstedeværelse eller fravær af stød.

Den fremgangsmåde har den åbenlyse ulempe at den kun er gearet til at blive brugt på optagelser som er tilrettelagt med henblik på dette konkrete formål fordi den slags matchede par nok sjældent vil forekomme ved et tilfælde i almindelig spontan tale. Fremgangsmåden vil kunne bruges til en indledende test af den opstillede hypotese, men vil sandsynligvis komme til kort i forhold til at undersøge stød i optagelser af tale generelt. Det vil sandsynligvis være forbundet med en betydelig usikkerhed at sige noget om eventuelle ændringer i stemmekvaliteten i forbindelse med stød i et konkret tilfælde hvis man ikke har et matchende modstykke at holde det op imod. Hvis hypotesen viser sig at holde til en indledende test, vil der derfor være anledning til at gå videre med forsøg på at isolere signalet fra kilden.

5.4.1 Et eksempel: Interaktion mellem $H1:H2$ og $F1$

Der er mange forskellige akustiske mål der kan relateres til aspekter ved stemmekvalitet. Maryn et al 2009 der er en metaanalyse af akustiske mål for "overall voice quality", inkluderer 87(!) forskellige akustiske markører. Det kan ses som udtryk for at der ikke er en etableret norm for hvordan stemmekvalitet bedst kvantificeres akustisk. I det følgende kapitel (6) vil jeg præsentere de metoder jeg har tænkt mig at benytte. Jeg vil give et enkelt eksempel her for at illustrere nogle af de udfordringer der er ved at kvantificere stemmekvalitet på baggrund af det direkte talesignal.

Et hyppigt anvendt mål er det relative styrkeforhold mellem den første og den anden deltone ($H1:H2$) der som nævnt varierer med graden af kompression og/eller med OQ. Problemet ved at beregne styrkeforholdet mellem $H1$ og $H2$ for output-spektret er at det vil være følsomt for hvor tæt $F1$ ligger på den første og især den anden deltone, dvs. $H2$. I de fleste tilfælde vil det være sådan at jo højere grundtonen er og jo lavere $F1$ er, desto mere vil den anden deltone ($H2$) blive forstærket af $F1$ -resonansen,

og som følge af det vil forholdet mellem H1 og H2 blive forskudt i retning af at indikere en mere komprimeret stemmekvalitet.

5.5 Periodicitet: yderligere udfordringer og muligheder

Som nævnt er det forventeligt at der vil forekomme passager med uregelmæssigheder i stemmelæbesvingningerne i forbindelse med stød. Uregelmæssige stemmelæbesvingninger udgør på forskellig måde en udfordring i forhold til at kvantificere stemmekvalitet.

5.5.1 Udfordring for bestemmelse af H1:H2

Uregelmæssigheder i periodiciteten vil give nogle vanskeligheder i forhold til en akustisk analyse baseret på overtonestrukturen. Det er kun meningsfyldt at tale om H1:H2 når signalet er så tilpas periodisk at der er en genkendelig overtonestruktur. For i praksis at kunne beregne styrkeforholdet er det desuden nødvendigt at kunne bestemme grundtonen med en passende præcision. Typisk vil det være sådan at jo mindre (ensartet) periodisk signalet er, desto større vanskeligheder vil der også være med at bestemme grundtonen.

Diplofoni udgør et andet problem i forhold til tolkningen af spektrale karakteristika som fx H1:H2. Dels er der det principelle spørgsmål som går på hvad der skal regnes for grundtone når et signal med en given frekvens moduleres med en frekvens som er den halve (se diskussionen i 2.4.1). Hvis vi tager udgangspunkt i at vi ønsker at bestemme grundtonen af grundsignalet, altså det underliggende umodulerede signal, så kræver det særlig opmærksomhed i forbindelse med automatiseret bestemmelse af f_0 -forløbet (som fx f_0 -tracking ved hjælp af Praat) for at undgå at modulationsfrekvensen fejlagtigt bliver registreret som grundtonen.

Men selv med den ønskede bestemmelse af grundtonen er der potentielle problemer afhængig af om der kun forekommer modulation af amplituden eller om der forekommer modulation af frekvensen. Hvis det kun er amplituden der moduleres, vil H1:H2 forholdet ikke blive påvirket fordi amplitudemodulation kun resulterer i ét sæt sidebånd omkring deltonerne som med en modulationsfrekvens på $f_0/2$ vil havne midt mellem grundsignalets deltoner. Hvis frekvensen også moduleres (dvs. timingen af de enkelte glottale pulser), så har det også konsekvenser for grundsignalets deltoner. Det skyldes at frekvensmodulation medfører en vifte af sidebånd omkring grundsignalets deltoner (som nævnt i 2.3), og med en modulationsfrekvens på $f_0/2$ vil halvdelen af dem falde sammen med grundsignalets deltoner og dermed også påvirke fx H1:H2-forholdet.

5.5.2 Alternativer til H1:H2

Eftersom H1:H2-målet forudsætter en passende høj grad af periodicitet, er det sandsynligt at det ikke vil kunne bruges til at sige noget om de spektrale forhold i selve kernen af et prototypisk stød, som netop er karakteriseret ved ikke-periodiske

stemmelæbesvingninger. For at få noget at vide om de spektrale forhold der er det nødvendigt at se på nogle mål som kan beregnes uanset om signalet er periodisk eller ej. To muligheder er Spectral Tilt og CoG.

Spectral Tilt, som angives i dB pr. oktav, er et mål for spektrets overordnet faldende eller stigende tendens. Glottispektret ved almindelig modal stemthed har som tidligere nævnt typisk et fald på 12 dB pr. oktav. Spectral tilt i outputspektret påvirkes også af talekanalens filterfunktion og vil derfor variere med vokalkvaliteten.

CoG, "Centre of Gravity", som angives i Hz, er en anden enkel måde at karakterisere et spektrum på ved at angive dets "balancepunkt" inden for et specificeret frekvensområde – sagt med andre ord finder man den frekvens der deler et givent frekvensområde op i to dele sådan at der er lige meget akustisk energi i de to delområder. Eftersom energien i stemmekildens spektrum aftager med forskellig stejlhed afhængig af graden af kompression (se 5.2), så vil CoG i det samlede output ændre sig i takt med dette – alt andet lige. Ved en mere komprimeret stemmeklang hvor spektret falder mindre stejlt, og hvor der således vil være mere energi til stede i den øvre del af spektret end der ville være ved almindelig eller underkomprimeret stemthed, vil det spektrale balancepunkt naturligvis også ligge ved en højere frekvens. CoG er også følsom for både ændringer i grundtone og formantfrekvenser; Højere f_0 og højere F1 vil begge føre til højere CoG-værdier. Både CoG og H1:H2 påvirkes altså af f_0 og F1, dog ikke på samme måde. Forskellene undersøges nærmere i afsnittene 6.3 til 6.6.

5.5.3 Muligheder for beskrivelse af graden af periodicitet

Der er flere måder at undersøge og kvantificere i hvilken grad et signal er periodisk.

Jitter og shimmer er to mål som især bruges inden for logopædien til at karakterisere en stemmes periodicitet for at diagnosticere evt. lidelser. Jitter er et mål for hvor stor fluktuation der er i periodetiden. Shimmer er et mål for hvor stor fluktuation der er i amplituden fra periode til periode. Der er dog et forhold som begrænser anvendeligheden af disse to mål i denne sammenhæng: Petersen 2001 har vist at man for at få pålidelige bestemmelse af jitter og shimmer er nødt til at se på minimum 5 perioder. Ved en grundtone på fx 100 Hz ville de betyde et tidsvindue på 50 ms hvilket er i overkanten for det aktuelle formål.

Et andet mål der kunne være relevant, er Harmonicity to Noise Ratio (HNR) der er et mål for hvor stor andelen af det harmoniske (dvs. periodiske) indhold i et signal er i forhold til det ikke-periodiske (dvs. støj).

Jeg vil desuden foreslå en simpel måde at kvantificere graden af diplofoni i lyd-signalet: Som tidligere beskrevet i 2.3 og 2.4.1 kan diplofoni betragtes som modulation med $f_0/2$ som resulterer i sidebånd der ligger midt imellem de deltonefrekvenser som grundtonen giver ophav til. Sidebåndene vil således have frekvenser der er hele multipla af $f_0/2$. Tentativt kunne disse benævnes $H\frac{1}{2}$, $H1\frac{1}{2}$, $H2\frac{1}{2}$ etc. Hvis grundtonen kan følges med tilstrækkelig præcision i passager med diplofoniske svingninger, vil

man kunne få en kvantificering af graden af diplofoni ved at se på styrkeforholdet mellem fx $H1$ og $H1\frac{1}{2}$ (som altså betegner en frekvens midt mellem $H1$ og $H2$). Den foreslåede metode uddybes nærmere i 6.7 og 6.9.

5.5.4 Aperiodicitet udfordrer modeller for kildens funktionsmåde

I korte træk kan man sige at de modeller for stemmekilden jeg har peget på, kun kan forklare nogle af de måder stemmelæberne kan komme i ikke-regelmæssige svingninger på. Det antages at de forhold der bevirker lavere OQ (dvs. højere kompression), i yderste konsekvens fører til at timingen af de glottale pulser bliver ustabil, se Stevens 2000, p. 85 (pp. 82-85).

Ud fra sådan en model må man forvente at overgangen til uregelmæssige svingninger eller diplofoni må forekomme i forlængelse af et stigende eller opadgående $H1:H2$ -forløb, og tilsvarende at overgangen fra uregelmæssige stemmelæbesvingninger til periodiske svingninger følges af et faldende $H1:H2$ -forløb (se figur 4.1).

Ud fra samme model kunne man også få den tanke at overgang til knirk vil forekomme ved en bestemt grænseværdi. I outputspektret vil en evt. grænseværdi selvfølgelig afhænge af interaktionen med $F1$, dvs. afhænge af både f_0 og $F1$, så der vil være nogle usikkerhedsmomenter i forhold til at undersøge det på baggrund af outputspektret.

Det er også tænkeligt at der kunne vise sig en vis asymmetri i overgangen (hysteres). I så fald vil det formodentlig være sådan at skiftet fra modal til knirk sker ved et lidt højere kompressionsniveau end overgangen fra knirk til modal.

Modeller for måden stemmekilden fungerer på, fokuserer på at modellere de enkelte pulsers form. Det antages (implicit) at de glottale pulser kommer i et jævnt periodisk forløb, og det er på den baggrund at kildespektret beregnes. Men afvigelser fra et jævnt periodisk forløb må påvirke spektret og (dermed) den opfattede klang. Dette aspekt falder desværre uden for de beskrivelser af kilden der trækkes på³⁶.

Det er derfor et åbent spørgsmål hvilke konsekvenser diplofoni eller (kaotisk) uregelmæssighed i timingen af de glottale pulser har for lydens spektrum. Det er klart at diplofoni vil medføre "ekstra" deltoner (ved hele multipla af modulationsfrekvensen, dvs. $f_0/2$) og at helt aperiodisk forekommende glottale pulser vil resultere i et spektrum uden egentlige deltoner, men om – og i givet fald hvordan – den overordnede spektrale fordeling påvirkes af variation i periodiciteten er ikke velbelyst. Emnet berøres meget tentativt i forhold til diplofoni i afsnit 6.20.

36 Der findes mere sofistikerede forslag til modeleringer af stemmekilden og dens spektrum end de modeller jeg har refereret til, fx Hanson et al. 2001, men tilsyneladende endnu ikke nogen der takler de akustiske aspekter af diplofoni og uregelmæssige svingninger. Hanson et al. 2001 diskuterer emnet, men nøjes med at konstatere at det har relevans: "In summary, irregular vibration should be included in models of phonation [...]" (p. 456).

Det kan også bemærkes at de nævnte modeller (Stevens 2000) ikke redegør for hvordan diplofoniske svingninger opstår. Modellerne kan heller ikke redegøre for hvordan de uregelmæssigheder i stemmelæbesvingningerne som beskrives i Slifka 2006, kan opstå.

5.5.5 Formulering af akustisk-fonetiske delhypoteser

Som nævnt i kapitlets begyndelse er den præsenterede hypotese om stød formuleret i fonetiske termer som ikke direkte giver muligheder for en empirisk afprøvning. For at kunne teste hypotesen med akustisk fonetiske metoder er man nødt til at overføre den til det domæne, og dermed reformulere hypotesen i akustisk fonetiske termer. På baggrund af de sammenhænge der er blevet diskuteret i løbet af kapitlet, kan man formulere fire delhypoteser som ekspliciterer nogle forhold og sammenhænge som man må forvente at kunne iagttage med akustiske metoder, hvis stød udfolder sig som et stemmekvalitetsforløb på den måde det er formuleret i hypotesen i afsnit 4.5.

De fire forhold man må forvente at finde, er:

- I. Et bestemt forløb for de akustiske mål der kan kobles til OQ.
- II. En bestemt sammenhæng mellem de akustiske mål der kan kobles til OQ og forekomsten af uregelmæssige stemmelæbesvingninger.
- III. En bestemt sammenhæng mellem intensitetsforløbet og forekomsten af uregelmæssige stemmelæbesvingninger.
- IV. En bestemt sammenhæng mellem intensitetsforløbet og de akustiske mål der kan kobles til OQ.

Ad I: Hvis stød udfolder sig som et stemmekvalitetsforløb på den måde det er formuleret i hypotesen, må man forvente at de akustiske mål der kan relateres til OQ – dvs. H1:H2 og måske andre mål som CoG eller spectral tilt – vil vise at OQ mindskes og øges igen som tegn på stigende-faldende ændring i kompressionen.

Ad II: I stød med uregelmæssige svingninger må man forvente at de akustiske mål der kan relateres til OQ, vil vise at OQ mindskes (som tegn på stigende kompression) op til passagen med uregelmæssige svingninger, og efterfølgende øges igen (som tegn på faldende kompression). I det omfang det kan lykkes at få indikationer af OQ i passager hvor stemmelæberne svinger uregelmæssigt, må det forventes at se tegn på lavere OQ (som indikation af at kompressionen er højere) sammenlignet med passager med regelmæssige svingninger.

Ad III: Det dyk i intensitetsforløbet som er karakteristisk for stød, indikerer hvornår den glottale konstriktion er kraftigst. I stød med uregelmæssige svingninger må man derfor ud fra Ladefogeds kontinuum forvente at intensitetsdykkets minimum findes i passagen med uregelmæssige svingninger.

Ad IV: Det dyk i intensitetsforløbet som er karakteristisk for stød, indikerer hvornår den glottale konstriktion er kraftigst. Ud fra Ladefogeds kontinuum må man forvente at OQ er lavest (som indikation af at kompressionen er højest) tæt ved det tidspunkt hvor intensitetsforløbet når sit minimum, hvis den tænkte ændring i stemmekvalitet udfolder sig i skalaens øvre del. Forbeholdet skyldes at der sandsynligvis ikke er tale om en monoton sammenhæng mellem OQ og intensitet over hele skalaen. Som det blev nævnt i afsnit 4.2, (s. 55) og som det fremgår af tabel 4.1, så må intensiteten af den lyd som stemmekilden genererer, formodes at være lavere for de modi der findes i begge ender af skalaen. I stød med uregelmæssige svingninger vil det formodentlig kun være muligt at efterprøve forventningen hvis det kan lykkes at få akustiske indikationer af OQ i de passager hvor stemmelæbernes svingningsmønster er uregelmæssigt.

Opsummerende kan man sige at punkt I overfører hovedtrækket i hypotesen til det akustisk fonetiske felt. Punkt II ekspliciterer sammenhængen mellem OQ og uregelmæssige svingninger ud fra forventningen om at de uregelmæssige svingninger der forekommer i forbindelse med stød, kan regnes for knirk. Punkterne III + IV ekspliciterer sammenhæng mellem intensitetsdykket og to forhold der relateres til knirk: uregelmæssige svingninger og lav OQ.

6 Akustisk analyse – implementering og test af metoder

I dette kapitel vil jeg forsøge at give et overblik over de akustiske metoder som vil blive benyttet i analysen. Der er hovedsagelig tale om almindelige akustiske metoder, men jeg vil introducere to nye analyser som kan betegnes som variationer over nogle af de øvrige metoder. De to nye metoder vil jeg i mangel af bedre navne kalde "Band Energy Difference" (BED) og " $H1:H1\frac{1}{2}$ ". Jeg vil begrunde hvorfor de enkelte metoder er valgt ud, og forklare tankerne bag de to alternative mål jeg har fundet på som supplement til de etablerede standardmål. I de tilfælde hvor der ikke er tale om standardværktøjer der kan tages direkte fra Praats værktøjskasse, beskrives det hvordan målene er implementeret. Der gives også en introduktion til den grafiske fremstilling af analyserne sådan at læseren forhåbentlig lettere kan orientere sig i de plots der genereres.

Det har været målet at konstruere et værktøj, konkret i form af et Praat-script, som kunne give en samling af praktisk anvendelige akustiske mål som kunne bidrage til en forståelse af de ændringer i stemmekvaliteten som forekommer i forbindelse med stød.

Det præsenterede værktøj skal ikke opfattes som den definitive fremgangsmåde eller implementering, men må ses som det første skridt i retning af at lave et værktøj med det specifikke mål at afprøve den opstillede hypotese. Samtidig vil det måske også kunne finde anvendelse til andre formål hvor det måtte være relevant at visualisere akustiske forhold der har med stemmekvalitet eller klang at gøre.

For at hypotesen kan testes er det nødvendigt at finde metoder som kan fortælle noget om relativt hurtige klangforandringer. Det er altså et krav til de akustiske analyser der skal benyttes, at de kan beregnes med et tilstrækkeligt kort tidsvindue.

Typiske vil lange vokaler have en udstrækning i tid på mellem 80 og 200 ms. Hvis det skal være muligt at følge de forandringer der måtte forekomme fx i løbet af den første halvdel af en lang vokal med stød, og hvis man fx som udgangspunkt tænker sig at det kræver fire målepunkter for at få en passende beskrivelse af forløbet i den del, betyder det at et tidsvindue³⁷ i omegnen af 10-20 ms vil være ønskeligt.

37 Akustiske træk kan kun opgøres for et tidsinterval, ofte benævnt analysens tidsvindue.

6.1 Oversigt over valgte metoder

På baggrund af de overvejelser jeg fremlagde i kapitel 5 vælger jeg at fokusere på overfladenære akustiske analyser. De akustiske mål jeg har valgt og som vil blive behandlet mere udførligt i det følgende, kan deles i fire kategorier:

1. Mål som beskriver forhold ved den spektrale (frekvensmæssige) sammensætning der kan relateres til stemmens (stemmekildens) klang:
 - $H1:H2$ (styrkeforholdet i outputspektret mellem H1 og H2)
 - CoG (Centre of Gravity)
 - BED^{38} (Band Energy Difference målt som forskellen i intensitet mellem området fra 0 Hz til $H1\frac{1}{2}$ og intensiteten i området fra $H1\frac{1}{2}$ to 1000 Hz)
 - $H1:H1\frac{1}{2}$ -se nedenfor
2. Mål for periodicitet:
 - $H1:H1\frac{1}{2}^{38}$ (Styrkeforholdet i outputspektret mellem H1 og et tilsvarende frekvensbånd mellem H1 og H2, her kaldet $H1\frac{1}{2}$. Energi i dette område kan tages som indikation for forekomst af diplofoniske svingninger)
 - HNR (Harmonicity to Noise Ratio: Styrkeforholdet mellem periodiske og aperiodiske komponenter i signalet)
3. Hjelpe-analyser som enten er nødvendige for beregningen af nogle af ovennævnte mål, eller som fungerer som referenceramme for de kurver der tegnes:
 - Oscillogram
 - Smalbåndsspektrogram (NB-spektrogram)
 - Bredbåndsspektrogram (WB-spektrogram)
 - f_0 – grundtone
4. Ekstra analyser som er mere indirekte relateret til afprøvningen af hypotesen:
 - Intensitet
 - Formantværdier

³⁸ BED og $H1:H1\frac{1}{2}$ forklares nærmere i de følgende afsnit 6.5 og 6.7.

6.2 Implementering og introduktion til den visuelle fremstilling

Beregning af de valgte mål foregår i ét Praat-script, se Appendix B-I. De centrale dele af scriptet som vedrører de egentlige målinger, er samlet i appendix B-II. Alle kurveforløb præsenteres i et samlet output, se figur 6.1. Den grafiske fremstilling er bygget op over tre standardanalyser der præsenteres i tre rammer:

- I Oscillogram
- II Smalbånds-spektrogram (Narrow Band – NB)³⁹
- III Bredbånds-spektrogram (Wide Band – WB)

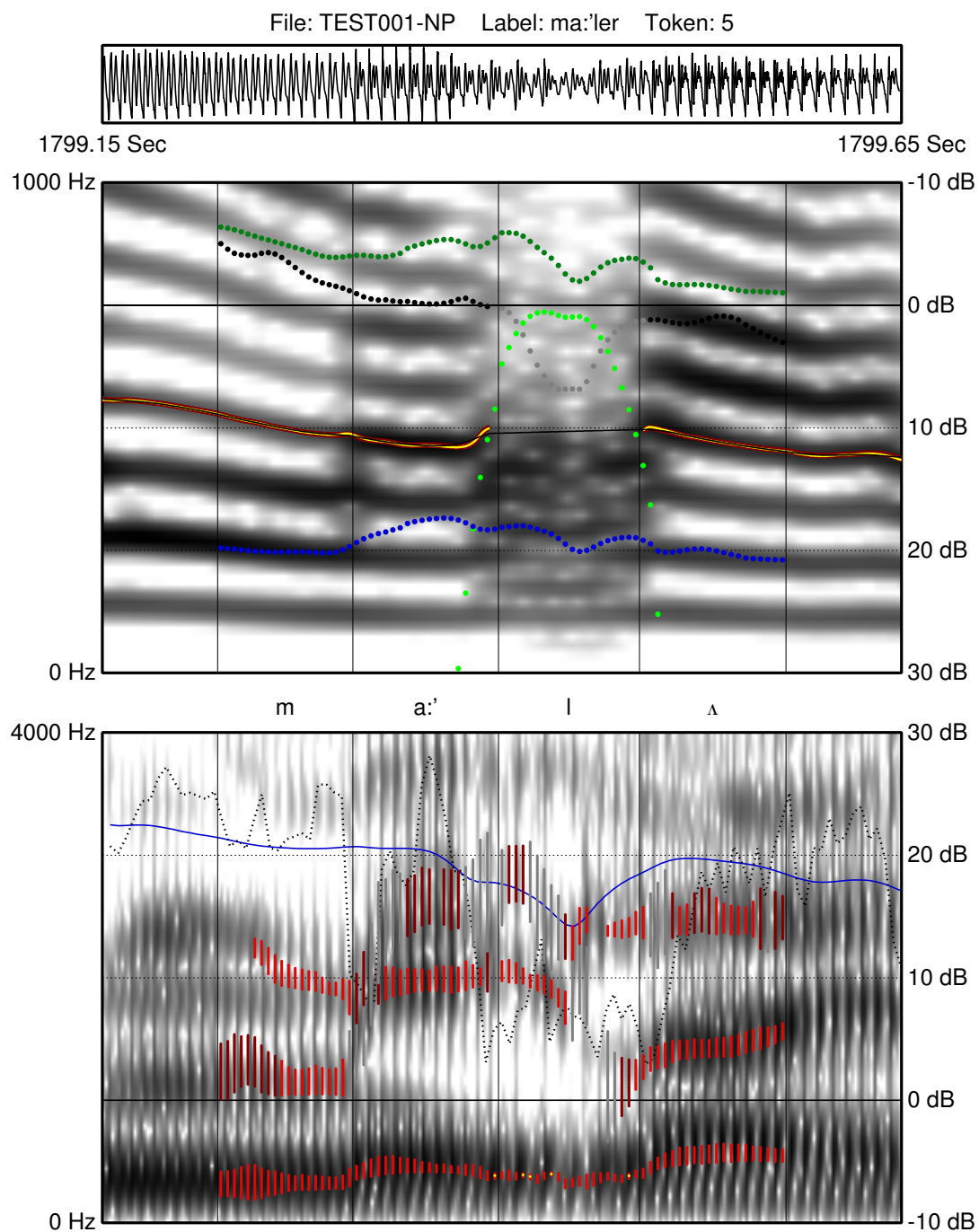
Tabel 6.1 på næste side giver en samlet oversigt over analyserne; hvordan de er placeret i de tre rammer, og hvilke farver der anvendes i plottene for de enkelte kurveforløb.

Smalbåndsspektrogrammet er beregnet med en vinduesbredde på 0.05 s. Sværtningsomfanget dækker et dynamikområde på 40 dB. For tydeligere at kunne se de enkelte deltoner plottes kun frekvensområdet fra 0 til 1000 Hz. Bredbåndsspektrogrammet er beregnet med en vinduesbredde på 0.005 s og dækker et dynamikområde på 35 dB og frekvensområdet fra 0 til 4000 Hz. De to spektrogrammer og oscillogrammet dækker samme tidsinterval og er nøjagtigt synkroniserede. De fungerer som baggrund og referenceramme for den grafiske fremstilling af de øvrige analyser.

39 Smalbåndsspektrogrammet (se figur 5.1, ramme II) giver en frekvensmæssig opløsning som gør at man kan adskille de enkelte deltoner (de vandrette mørke bånd i ramme II), men medfører samtidig en begrænset tidslig opløsning. Bredbåndsspektrogrammet (se figur 5.1, ramme III) har en bedre tidslig opløsning sådan at det ofte bliver muligt at se de enkelte glottale pulser (de lodrette mørke bånd i ramme III), men på bekostning af den frekvensmæssige opløsning.

Akustiske analyser			
	Analyse	Farver	Beskrivelse
I	Oscillogram	Sort	
II	NB-spektrogram	Sort/hvid	0-1000 Hz, 35 dB
	f_0	Brun, gul, sort	Grundtonen, to estimator + smoothing og interpolation
	H1:H2	Sort	Styrkeforholdet mellem H1 og H2
	CoG	Blå	Centre of Gravity for frekvensområdet 0-500 Hz
	BED	Mørk grøn	Band Energy Difference for 0 – H1½ Hz relativ til H1½ – 500 Hz, hvor $H1\frac{1}{2} = f_0 * 1.5$
	H1:H1½	Lys grøn	Styrkeforholdet mellem H1 og H1½, dvs. området mellem H1 og H2
III	WB-spektrogram	Sort/hvid	0-4000 Hz, 40 dB
	HNR	Sort stiplede	Harmonicity to Noise Ratio
	Intensitet	Blå	Lydsignalets styrke eller energiindhold
	Formanter	Gul, rød, brun, grå	F1, F2, F3 + båndbredder

Tabel 6.1. Oversigt over analysernes placering i de tre rammer og de farver der anvendes.



Figur 6.1. Eksempel på output fra målescriptet. Opbygningen gennemgås i de følgende afsnit (6.3 til 6.12)

6.3 H1:H2

Under i øvrigt sammenlignelige forhold ændrer styrkeforholdet mellem grundtonen og første overtone sig som beskrevet i foregående kapitel med stemmekildens kompressionsgrad. H1:H2 kan bestemmes ved at måle det relative energiindhold i to smalle frekvensbånd centreret omkring henholdsvis grundtonen og første overtone. I det script jeg har skrevet, sker det ved at benytte funktionen Band Energy Difference. De nærmere detaljer om hvordan grundtonen (H1 aka f_0) fastlægges, beskrives i afsnit 6.10. Frekvensen for den anden deltone (H2) beregnes ud fra f_0 -estimatet ved at gange med 2.

Hvis f_0 -estimerne gav fuldstændig nøjagtige resultater, og hvis signalet fra stemmekilden var fuldstændig periodisk, så ville man i princippet få den mest nøjagtige bestemmelse af deltonernes styrkeforhold og det mindste bidrag fra uvedkommende støjkomponenter ved at måle i så smalle filterbånd som muligt. Men da f_0 -estimatet ikke kan være fuldstændigt nøjagtigt, bl.a. fordi signalet fra stemmekilden aldrig vil være fuldstændig periodisk, så må filterbåndbredderne være brede nok til at indfange deltonerne på trods af unøjagtigheden. Som nævnt beregnes H2 ud fra f_0 -estimatet ved at gange med to. Bestemmelsen af H2 vil derfor nødvendigvis være behæftet med dobbelt så stor unøjagtighed som unøjagtigheden på f_0 -estimatet.

Jeg ved ikke af nogen måde hvorpå man med sikkerhed kan afgøre nøjagtig hvilke filterbåndbredder der vil være de optimale. De konkrete filterbåndbredder på 20 Hz er valgt så de skønsmæssigt giver en passende balance mellem følsomhed for støj og følsomhed for unøjagtigheder i f_0 -estimatet. Jeg vender tilbage til en test af beregningen af H1:H2 i afsnit 6.14.

Jeg har valgt at beregne H1:H2 på baggrund af det udglattede grundtoneestimat (plottet med fuldt optrukket sort linje, se i øvrigt videre afsnit 6.10). Gevinsten ved udglatningen er hovedsagelig at bidraget fra kortvarige glitches i f_0 -bestemmelsen mindskes. I de tilfælde opnås en mere nøjagtig bestemmelse af deltonernes frekvenser. Ulempen er at hurtige, men typisk svage fluktuationer i grundtonen, jævnes ud. I de tilfælde medfører det en lidt mindre præcis bestemmelse af f_0 og dermed overtonerne. Det er mit indtryk at der netto er en gevinst ved at basere H1:H2 på et udglattet grundtoneestimat, men fastlæggelsen af H1 og H2 og målingen af deres indbyrdes styrkeforhold vil nok med fordel kunne forfines i en eventuel fremtidig implementering.

H1:H2-forløbet plottes med sorte prikker i ramme II (smalbåndsspektrogrammet). Når bestemmelsen af f_0 er mindre sikker, plottes H1:H2 med grå. f_0 -bedømmelsen betragtes som mindre sikker når HNR er under -10 dB, eller hvis forskellen mellem de to f_0 -estimer overstiger 5 %, og dermed også de passager hvor der kun er et interpoleret bud på f_0 . Bemærk, disse kriterier siger ikke noget om hvorvidt det konkrete grundtoneestimat er tilstrækkeligt præcist til at H1:H2-forholdet kan måles.

Det giver kun en indikation af at der er forhold som gør at der er grund til at være særligt opmærksom på om analyserne ser troværdige ud.

Styrkeforholdet mellem H1 og H2 udtrykkes pr. tradition med udgangspunkt i H1. For at få overensstemmelse mellem de forskellige kurveforløb er dB-skalaen (markeret til højre for ramme II) vendt om sådan at +30 dB er for neden og -10 dB for oven. H1:H2-kurveforløb som bevæger sig opad, indikerer således en ændring i retning af mere komprimeret stemmekvalitet, med mere energi ved H2, mens modsat lavere liggende kurveforløb indikerer en mindre komprimeret stemmekvalitet. Den fuldt optrukne vandrette linie 3/4 oppe i spektrogrammet markerer nullinien, dvs. hvor H1 og H2 har samme styrke.

6.4 CoG (Centre of Gravity)

Centre of Gravity som akustisk mål er en analogi til den mekaniske fysiks massemidt-punkt. I denne sammenhæng betegner det balancepunktet for et givet lydsignals frekvensspektrum, beregnet for et nærmere bestemt frekvensområde og anskuet på en lineær frekvensskala. Alt andet lige vil et lydsignal med mest energi ved højere frekvenser have en højere CoG-værdi end et lydsignal med mest energi ved lavere frekvenser.

Signalet fra stemmekilden dominerer den nedre del af frekvensspektret. Støjlyde fra obstruenter (som lukkelyde eller frikativer) kan dominere højere oppe i frekvensspektret. Overgangen til og fra især frikativer kan være ret langsomt glidende. For så vidt muligt at undgå at eventuelle støjlyde fra konsonantiske nabosegmenter påvirker målingen af de klangforandringer som har med stemmekilden at gøre, beregnes CoG derfor for frekvensområdet 0-500 Hz hvor energien fra det periodiske signal formodes at dominere. CoG plottes med blå prikker i ramme II (smalbåndsspektrogrammet), se figur 6.1.

CoG beregnes ved hjælp af Praats indbyggede algoritme (spectrogram > to spectrum (slice) > get centre of gravity) på baggrund af et spektrogram beregnet med en vinduesbredde på 0.06 s, dvs. med en anelse smallere bånd end det som bruges som baggrund i ramme II, og med en øvre grænsefrekvens på 500 Hz – se videre appendix B.

6.5 BED (Band Energy Difference)

Hvor H1:H2 kun kan bestemmes når signalet er tilstrækkeligt periodisk, så kan CoG bestemmes uanset om signalet er periodisk eller ej. Begge mål påvirkes af f_0 og formantstrukturen, men på forskellig måde: Højere f_0 og højere formantfrekvenser resulterer begge i højere CoG. For H1:H2 er billedet lidt mere komplekst, men interaktionen vil være størst jo højere grundtonen er, og jo lavere F1 er, og typisk vil højere f_0 og lavere F1 medføre relativt mere energi ved H2 end ved H1.

Band Energy Difference er et forsøg på at skabe et mål som er en mellemtang mellem de to nævnte mål, både hvad angår afhængighed af periodicitet og præcis bestem-

melse af grundtonen, og hvad angår interaktionen med grundtonehøjden og formantstrukturen.

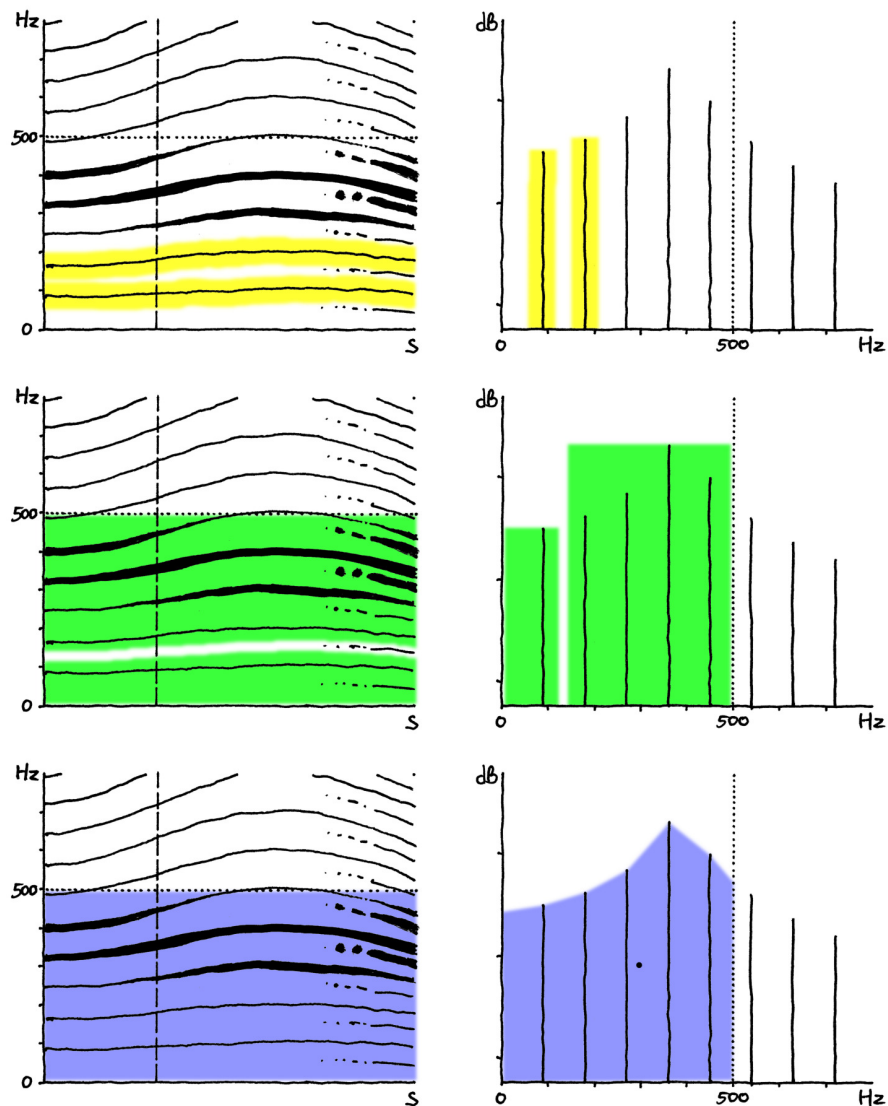
Grundideen er at lave en grovere udgave af $H1:H2$ som samtidig tager et lidt større frekvensområde i betragtning. Metoden går ud på at dele et valgt frekvensbånd, her 0 til 500 Hz, i to dele midt mellem grundtonen og første overtone og sammenligne energiindholdet i de to dele af frekvensbåndet. Ved at regne på frekvensbånd som er noget bredere end de som indgår i beregningen af $H1:H2$, slækkes kravet til nøjagtigheden af bestemmelsen af f_0 noget. Det er dog stadig vigtigt at grænsen mellem de to frekvensbånd kommer til at ligge mellem $H1$ og $H2$.

For at kunne beregne BED i de passager som ikke er tilstrækkeligt periodiske til at der kan bestemmes en grundtone (enten fordi de er ustemte, eller fordi stemmelæbesvingningerne er uregelmæssige), benyttes en interpoleret og let udglattet f_0 . I de interpolerede passager må BED betragtes som enten et kvalificeret gæt eller med en mere abstrakt reference til de omgivende passager som er periodiske. Fx vil knirkede passager som er omgivet af passager med en lav grundtone, resultere i andre BED-værdier end hvis den samme passage havde været omgivet af passager med en højere grundtone.

Band Energy Difference-målet plottes med grøn farve i ramme II (smalbåndsspektrogrammet), se figur 6.1. Kurveforløbet bevæger sig opad når der er mere energi i det øvre frekvensbånd, og indikerer dermed en mere komprimeret stemmekvalitet.

6.6 Motivation for $H1:H2$, CoG og BED

Pointen med de tre relaterede mål er at kunne kvalificere eventuelle udsagn om variationer i stemmekvaliteten. Relevansen af målene hviler på den antagelse at lyd fra stemmekilden med relativt mere energi ved frekvenser over grundtonen vil blive opfattet som mere komprimeret. I den forstand er der tale om en operationalisering af et væsentligt aspekt ved begrebet kompression. De tre metoder udviser forskellig følsomhed for interaktionen mellem deltonerne og formantstrukturen, og samtidig forskellig følsomhed for støjkomponenter og i hvor høj grad signalet er periodisk. I det omfang de tre mål giver overensstemmende indikationer, vil det styrke analysen. I de tilfælde hvor der ikke er overensstemmelse mellem de tre mål, vil man få en indikation af hvor målene er mindre pålidelige. Det er samtidig tænkeligt at man på grund af de forskellige interaktionsmønstre vil kunne regne ud hvilke af de tre mål som påvirkes mindst af de forstyrrende faktorer i et givet tilfælde. Figur 6.2 giver en grafisk oversigt over de tre metoder.



Figur 6.2. Illustration af H1:H2, BED og CoG. Til venstre er illustreret et stiliseret spektrogram og til højre en stiliseret spektral-analyse. De farvede områder illustrerer de frekvensbånd som tages i betragtning. For H1:H2 og til dels BED følger de deltonernes forløb. For BED gælder det kun grænsen imellem de to frekvensområder. Den nedre frekvensgrænse på 0 Hz og den øvre grænse på 500 Hz (markeret med en stiplet linje) ligger fast. Grænserne for det frekvensområde der tages i betragtning ved beregningen af CoG ligger også fast på 0-500 Hz.

6.7 H1:H1½ – et mål for diplofoni

Diplofoni kan, som der blev argumenteret for i afsnit 2.4.1, tolkes som modulation (frekvens og/eller amplitudemodulation) med en frekvens der er den halve af grundtonen. Ved diplofoniske svingninger vil der derfor forekomme energi ved frekvenser der ligger midt mellem deltonerne. Disse må betegnes som sidebånd, se figur 6.3 th. Ved at måle styrkeforholdet mellem to smalle frekvensbånd centreret omkring henholdsvis grundtonen og det sidebånd der ligger mellem første og anden deltone, dvs. ved $1.5 \times$ grundtonen, kan man få et mål for graden af diplofoni. Da det sidebånd jeg vil måle på, netop har en frekvens som er $1.5 \times$ grundtonen, vælger jeg at betegne det H1½ – eller for at være helt præcis: med H1½ refererer jeg til den frekvens mellem H1 og H2 hvor der vil opstå energi hvis et signal moduleres med en frekvens som er den halve af signalets egen frekvens⁴⁰.

Da der ofte kan være vanskeligheder med at bestemme grundtonen i passager med diplofoni, er målet baseret på en interpoleret og udglattet f_0 -kurve (på samme måde som BED-målet, se ovenfor). Ved tolkning af H1:H1½- målet er det derfor afgørende løbende at vurdere om det interpolerede grundtoneforløb er troværdigt.

Forløbet af de limegrønne prikker indikerer hvor meget energi der er ved det sidebånd der ligger mellem H1 og H2 i forhold til grundtonen (H1). Højereliggende værdier indikerer mere energi ved H1½, dvs. som tegn på diplofoni. Når signalet er jævnt periodisk, vil der være meget lidt energi i området midt mellem H1 og H2 – ofte langt uden for de værdier der plottes. Ved tydelig diplofoni kan styrken af H1½ komme op i samme størrelsesorden som H1. I figur 6.1 ses tydelig diplofoni i [l]'et.

6.8 Harmonicity to Noise Ratio (HNR)

Harmonicity to Noise Ratio (HNR) er et mål som angiver styrkeforholdet i signalet mellem den del af energien i spektret som kan tilskrives periodiske komponenter (dvs. energi ved deltonerne), og den del som må tilskrives støjkomponenter (dvs. energi mellem deltonerne). HNR-målet tjener to formål i denne sammenhæng.

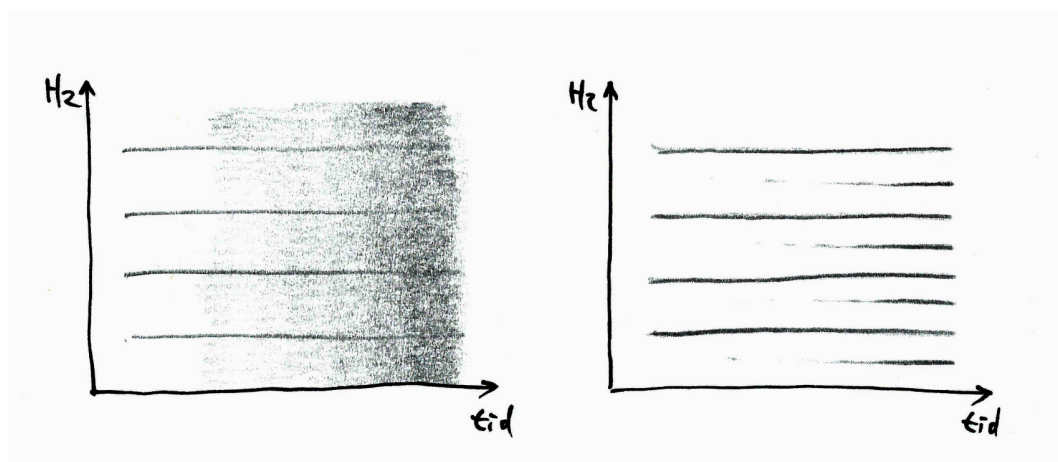
Det ene er at det bruges som en indikator for pålideligheden af f_0 -estimeringen. Når HNR-målet er højt, er der stor sandsynlighed for at f_0 -algoritmen vil præstere et tilstrækkelig præcist resultat. Når HNR målet falder, er der øget risiko for at få et upålideligt f_0 -estimat. Det stemmer overens med at det samtidig bliver mere problematisk overhovedet at tale om en grundtone. I forbindelse med f_0 -estimeringen er jeg nået frem til at bruge -10 dB HNR som grænseværdi, dvs. at når HNR-værdien er under dette, forkastes f_0 -estimererne. Denne grænseværdi begrundes således: Når HNR er over +10, er der sjældent problemer med f_0 -estimeringen. Når HNR falder under 0 dB i længere passager, sker det ofte ganske brat og til niveauer langt under -10 dB. I disse passager kommer f_0 -algoritmen som regel ikke med noget bud på f_0 , men når det sker er sandsynligheden for at det er en brugbar værdi, meget lille. Fordi

⁴⁰ NB, H1½ er en ad hoc betegnelse, og ikke etableret terminologi!

ændringen i HNR omkring længere passager med lave værdier sker ganske brat, vil en ændring af grænseværdien i intervallet mellem -10 dB og 0 dB ikke have nogen nævneværdig betydning for disse passager, men vil først og fremmest have en indflydelse på om kortvarige dyk i HNR skal føre til forkastelse af f_0 -estimatet. Med den fastsatte grænseværdi tjener HNR-målet til at forkaste f_0 -estimer som med overvejende sandsynlighed er upålidelige. (Man skal være opmærksom på at de værdier jeg nævner her, ikke ukritisk vil kunne overføres til andre optagelser, især ikke hvis disse indeholder støjkomponenter som ikke hører til talesignalet). Ud over den validerende funktion er HNR-målet samtidig et bud på en kvantificering af den uregelmæssighed i stemmelæbesvingningerne som ofte optræder i forbindelse med stød. HNR beregnes ved hjælp af Praats funktion *Harmonicity(cc)*, med indstillingerne 0.005 50 0.1 1, og plottes med sort stiplede linie i ramme III (bredbåndsspektrogrammet).

6.9 Motivation for HNR og $H1:H1\frac{1}{2}$

HNR er følsom for alle forhold der influerer på hvor klar overtonestrukturen er, uanset om det skyldes at den periodiske komponent i signalet bliver mindre periodisk, eller om det skyldes tilføjelse af anden støj (se figur 6.3). HNR vil altså reagere både på ændringer i periodiciteten, dvs. fx knirk, og på frikativ støj i talesignalet eller andre støjklender (pust i mikrofonen, båndstøj, trafik, etc.).



Figur 6.3. Sammenligning af stiliseret spektrogram for periodisk signal med overtoner hvor der (tv.) gradvis tilføjes støj, eller (th.) gradvis indføres diplofoni, som medfører sidebånd forårsaget af modulationen.

$H1:H1\frac{1}{2}$ er designet til at give udslag ved en bestemt form for "regelmæssig uregelmæssighed", nemlig diplofoni, men den vil også give udslag hvis energiniveauet i de undersøgte frekvensbånd omkring $H1$ og midterområdet mellem $H1$ og $H2$ ændres af andre årsager, fx på grund af støj.

Tilstedeværelse af støj (se figur 6.3 tv.) vil påvirke både HNR og $H1:H1\frac{1}{2}$ -målet.

Diplofoni (se figur 6.3 th.) vil give sig udslag i et ændret $H1:H1\frac{1}{2}$ -forhold, men bør ikke påvirke HNR. Diplofoni (eller, teknisk set, modulation med en frekvens der er den halve af grundtonen) gør ikke i sig selv signalet mindre periodisk, men medfører at periodetiden – set fra en strengt teoretisk synsvinkel – ændres til den periodetid modulationen har. Hvordan $H1:H1\frac{1}{2}$ og HNR i praksis reagerer på diplofoni, afprøves i afsnit 6.19.

HNR kan beregnes uanset om signalet er periodisk eller ej. $H1:H1\frac{1}{2}$ kan kun meningsfuldt beregnes når signalet er tilstrækkeligt periodisk til at det giver mening at tale om en overtonestruktur. I praksis afhænger det først og fremmest af om grundtonen kan bestemmes. En sammenligning af $H1:H1\frac{1}{2}$ og HNR vil kunne bruges til at afgøre om en stigning i $H1:H1\frac{1}{2}$ kan tolkes som tegn på diplofoni eller ej.

6.10 f_0 -tracking

Pålideligheden af beregningerne af $H1:H2$, $H1:H1\frac{1}{2}$ og BED er afhængig af en tilstrækkelig præcis bestemmelse af grundtonen. De uregelmæssigheder i periodiciteten som ofte forekommer i forbindelse med stød, gør det notorisk vanskeligt at bestemme grundtonen. Der gøres derfor tre tiltag for at sikre en så pålidelig bestemmelse af grundtonen som muligt:

- Resultatet af to f_0 -analyser, udført på to lidt forskellige måder, sammenlignes. Hvis forskellen overstiger 5 %, forkastes f_0 -estimerne.
- Måleværdier fra passager med lav periodicitet vurderet ud fra måling af HNR forkastes.
- Visuel verifikation.

6.10.1 To f_0 estimer

Estimering af grundtonen (f_0) udføres ved hjælp af algoritmer som leder efter gentagelser (periodicitet) i signalet. Algoritmerne styres blandt andet med nogle parametre som begrænser udfaldsrummet. Fx vil man altid specificere det frekvensområde man forventer at grundtonen skal findes inden for.

Det er min erfaring at det resultat der kommer ud af en f_0 -algoritme, ofte enten er ganske præcist eller rammer betydeligt ved siden af. Justering af grundtonealgoritmens styringsparametre har som regel den største effekt i problematiske passager, og ved at justere på algoritmens parametre kan man ofte, om man så må sige, bringe den "i" eller "ud af fokus". Ved at sammenligne resultaterne fra to analyser lavet med forskellige parametre eller forudsætninger er det muligt at få en indikation af resultatets pålidelighed: Jo tættere de to resultater kommer på hinanden, jo mere sandsynligt er det at det er et pålideligt mål.

Gennem forsøg er jeg nået frem til at følgende to fremgangsmåder kompletterer hinanden godt:

- Et f_0 -estimat beregnet med en øvre grænse for f_0 -estimatet på 600 Hz.
- Et andet estimat beregnet med en øvre grænse på 300 Hz, men baseret på en version af lyden som er lavpasfiltreret ved 1000 Hz, implementeret ved en båndpasfiltrering med grænseværdierne 50-1000 Hz og en smoothing (overgangs-zone) på 100 Hz.

De øvrige instillinger er ens for de to estimater, se videre appendix B.

Rationalet bag fremgangsmåden er at den ene procedure i potentielt problematiske passager vil tendere til at komme med et estimat som ligger for højt, hvorimod den anden procedure har større sandsynlighed for at gætte for lavt. Estimatet godtages hvis de to værdier ikke skiller sig mere end 5 % fra hinanden. I den anvendte implementering af målescriptet bruges resultatet fra det lavpasfiltrede signal som udgangspunkt for de videre beregninger. Hvis forskellen er større end 5 %, markeres estimatet som mindre sikkert.

6.10.2 Visuel verifikation af f_0

Som tidligere nævnt har periodiske signaler en deltonestruktur hvor der er energi ved frekvenser som er hele multipla af grundtonen. I et smalbåndsspektrogram træder deltonerne frem som smalle mørke bånd (se figur 6.1, ramme II). Deltonerne afspejler nødvendigvis periodiciteten, og det nederste bånd svarer til grundtonen. Grundtonen kan ikke (automatisk) aflæses ud fra spektrogrammet, men man kan få en grov visuel verifikation af et givent f_0 -estimat ved at plote det oven i smalbåndsspektrogrammet.

Ændringer i grundtonen giver anledning til større ændringer i en given deltones placering, jo højere deltone man ser på: Eftersom fx den fjerde deltone har fire gange så høj frekvens som grundtonen, vil en ændring af grundtonen modsvares af en fire gange så stor ændring af den fjerde deltones frekvens. Ændringer i grundtonen ses altså lettere højere oppe i deltonerækken.

Ved at plote f_0 -estimatet med en frekvensskala som er 1/4 af det frekvensområde som smalbåndsspektrogrammet dækker, vil grundtonekurven (ideelt set) komme til at ligge midt i den fjerde overtone. På den måde vil man kunne vurdere overensstemmelsen mellem f_0 -estimatet og overtonestrukturen med fire gange så stor nøjagtighed som hvis man havde plottet f_0 -estimatet oven i den nederste deltone (aka grundtonen).

Der er ikke noget helligt ved det forhold på 1:4 jeg har valgt, men det rammer en rimelig balance mellem to modstridende forhold: På den ene side set er deltonernes udsving større, jo højere op i deltonerækken man går. På den anden side set er det sådan at deltonerne generelt bliver svagere og derfor vanskeligere at se, jo højere man kommer op i deltonerækken. Det sidste afhænger bl.a. af hvor komprimeret stemmen er. Ved forholdet 1:4 fungerer deltonerne som en løbende relativ skala for afvigelser

på denne måde: Falder grundtonekurven sammen med den tredje eller den femte deltone, betyder det at estimatet er henholdsvis +25 % eller -25 % ved siden af. Falder kurven midt mellem den fjerde deltone og en af de nærmeste deltoner, er afvigelsen omkring 12.5 %.

Smalbåndsspektrogrammet illustrerer også graden af periodicitet: Jo mindre (jævnt) periodisk signalet er, jo mere utydelig bliver overtonestrukturen også. HNR-målet er en kvantificering af det forhold.

f_0 -estimatet som er baseret på det lavpasfiltrerede signal, er plottet med gul, mens det andet f_0 -estimat er plottet med brun, se figur 6.1, ramme II.

6.10.3 Udglatning og interpolation

Til brug for udregningen af de mål der afhænger af f_0 , beregnes en let udglattet og interpoleret f_0 baseret på det lavpasfiltrerede signal. Det udglattede og interpolerede f_0 -forløb plottes med sort i ramme II.

6.11 Formanter

Fischer-Jørgensens undersøgelser giver modstridende indikationer af om vokaler med og uden stød har helt samme vokalkvalitet. På den ene side iagttager Fischer-Jørgensen en forskel mellem palatogrammer af vokalen [i:] sagt med eller uden stød. Palatogrammerne viser næsten enstemmende tegn på en snævrere vokalartikulation i de tilfælde hvor vokalen har stød. På den anden side viser de spektrografiske analyser ikke tegn på at formantværdierne i vokaler med og uden stød er forskellige.

Selvom de nævnte palatogrammer tyder på en snævrere artikulation af vokalen [i] når der er stød, så er forskellen ikke så iøjefaldende at nogen har bemærket at vokaler med stød generelt skulle have en anden vokalkvalitet end tilsvarende vokaler uden stød. Man må derfor formode at der må være tale om beskedne forskelle. Det er dog muligt at en evt. minimalt ændret vokalkvalitet ville kunne bidrage til at opfatte skellet mellem vokaler med og uden stød uden at vi opfatter det som en selvstændig forskel i vokalkvalitet.

En anden grund til at se på formantværdierne er at ændringen af glottis' konfiguration i forbindelse med stød kunne tænkes at have en lille indflydelse på formantstrukturen ved at det akustiske tab ved glottis påvirkes af den glottale konstriktion. Det kunne give anledning til forskelle i formanternes båndbredder.

Den stemmekvalitets-hypotese jeg har foreslået, giver ikke direkte anledning til at se på formantværdier, men hvis den glottale konstriktion der forekommer i forbindelse med stød, giver sig udslag i formanternes båndbredder, og hvis resultaterne bekræfter min hypotese, så kunne det give anledning til overvejelser om hvorvidt der kunne være en mere generel sammenhæng mellem stemmekvalitet og formantbåndbredder.

Formanterne estimeres ved hjælp af Formant(Burg) metoden (se appendix B for specifikke settings). Fordi jeg har en formodning om at der kan forekomme relevante forandringer i formanternes estimerede båndbredde i forbindelse med stød, har jeg fundet på en måde at illustrere båndbredderne sammen med formantværdierne.

Jeg har dels eksperimenteret med farvekodning, og dels med at illustrere båndbredden netop som et bånd med varierende bredde. I den aktuelle implementering har jeg valgt at kombinere de to metoder; se figur 6.1 ramme III. Farvekodningen afhænger af forholdet mellem båndbredden og formantværdien, efter følgende skema:

Båndbredde i forhold til formantfrekvens	farve
under 1/10	gul
mellem 1/10 og 1/5	rød
mellem 1/5 og 1/2	brun
mellem 1/2 og 1/1	grå

Båndbredden illustreres desuden som et bånd hvis bredde i Hz er den samme som den estimerede båndbredde, og som er centreret omkring den estimerede formantfrekvens. Hvis båndbredden for en given formant overstiger 1/1 af formantfrekvensen, plottes estimatet ikke – dels fordi båndbredden potentielt kan blive så stor at en formant ville dække hele spektret – dels fordi estimerer med meget stor båndbredde sjældent repræsenterer et pålideligt estimat.

6.12 Intensitet

Måling af intensitet er nødvendig for at kunne afprøve delhypotese III og IV jævnfør afsnit 5.5.5. Intensiteten måles med Praats Intensity-funktion og plottes med blå i ramme III. De specifikke settings fremgår af appendix B.

6.13 Afprøvning af de akustiske analyser på testmateriale

For at kunne afprøve analysemetoderne har jeg lavet nogle testoptagelser af lange vokallyde hvor jeg forsøgsvis varierer en fonetisk variabel ad gangen. Hensigten er få mulighed for at se om analyserne reagerer som forventet på variation i stemmens kompression og for at få en indikation af i hvilket omfang analysemetoderne påvirkes af faktorer som vokalkvalitet og grundtonehøjde. Det er samtidig væsentligt at se på i hvor høj grad H1:H2, BED og CoG giver samstemmende resultater.

Der er otte testoptagelser, fordelt på fire serier:

1. To optagelser hvor stemmekvaliteten varieres gradvist fra underkomprimeret (og let luftfyldt) over modal til overkomprimeret, mens grundtone og vokalkvalitet, som er henholdsvis [i] og [a], så vidt mulig holdes konstant. Lydoptagelserne findes som appendix A.I.1 og A.I.2.
2. To optagelser hvor grundtonen varieres fra lav til høj, mens kompressionen og vokalkvalitet så vidt muligt fastholdes på hhv. middel niveau (dvs. modal) og de to vokalkvaliteter [i] og [a]. Lydoptagelserne findes som appendix A.II.1 og A.II.2.
3. Tre optagelser med varierende vokalkvaliteter, middel grundtone, middel kompression. Første optagelse rummer urundede fortungevokaler fra [i] til [a], anden optagelse rummer bagtungevokaler fra [u] til [ɔ] og videre til [ɑ] (som strengt taget hverken er en for- eller bagtungevokal, men midt imellem) og endelig en serie snævre vokaler [i y u] der skifter fra urundet til rundet og dernæst fra for- til bagtunge. Lydoptagelserne findes som appendix A.III.1, A.III.2 og A.III.3.
4. En optagelse hvor der veksles mellem modal stemthed, knirk og diplofoni. Lydoptagelsen findes som appendix A.IV.1.

De to vokalkvaliteter [i] og [a] er valgt fordi de har meget forskellige F1-værdier. Ved [i] ligger den laveste formant F1 omkring 250-300 Hz. Alt afhængig af grundtonens frekvens vil H3, H2 eller evt. H1 være den deltone der forstærkes mest af resonansen ved F1. Ændringer i grundtonen vil derfor medføre betydelige ændringer i styrkeforholdet mellem H1 og H2. Ved [a] hvor F1 ligger oppe omkring 650-800 Hz, vil det være deltoner længere oppe i rækken der falder sammen med F1 – typisk H4-H6. Ændringer i grundtonen vil derfor ikke i samme grad medføre ændringer i styrkeforholdet mellem H1 og H2.

6.14 Test af filterbreddens indflydelse ved beregningen H1:H2

Som nævnt i forbindelse med beskrivelsen af H1:H2-målet i afsnit 6.3 er det ikke indlysende hvor brede eller smalle frekvensbånd det er mest hensigtsmæssigt at tage i betragtning ved beregningen af H1:H2-forholdet. Meget smalle frekvensbånd vil kræve en meget præcis bestemmelse af f_0 , og med bredere frekvensbånd løber man en øget risiko for at lyd der ikke har med H1 eller H2 at gøre, kommer til at influere på

resultatet. Jeg har som nævnt (i afsnit 6.3) skønnet at et frekvensbånd på 20 Hz udgjorde et passende kompromis, men jeg vil nu i det følgende illustrere at det har en relativt beskedne indflydelse på H1:H2-målet at ændre på bredden af de frekvensbånd der tages i betragtning.

Appendix C.I til C.VI indeholder eksempler på beregning af H1:H2 på baggrund af mere eller mindre snævre frekvensbånd centreret omkring estimer af H1 og H2 beregnet ud fra f_0 . Der er lavet testkørsler på testoptagelserne⁴¹ i serie 1, 2 og 4 samt den første optagelse i serie 3. Der er beregnet fem analyser med frekvensbånd varierende i bredde fra 5 -100 Hz, plottet med hver sin farve på denne måde:

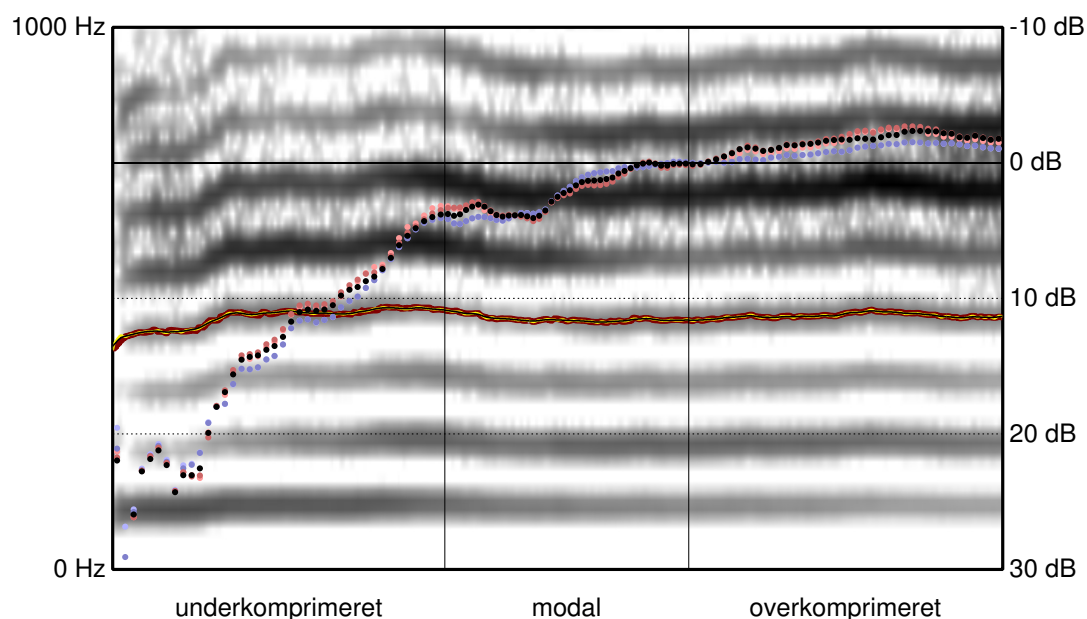
Farve	Frekvensbåndets bredde
lys blålig	=100 Hz
blålig	= 50 Hz
sort	= 20 Hz (Denne værdi benyttes i de videre analyser)
rødlig	= 10 Hz
lys rødlig	= 5 Hz

Af pladshensyn nøjes jeg i det følgende med at gengive ramme II hvor H1:H2, BED og H1:H1½ plottes undtagen når der er behov for at referere til ramme I eller III. De komplette figurer findes i appendix C.

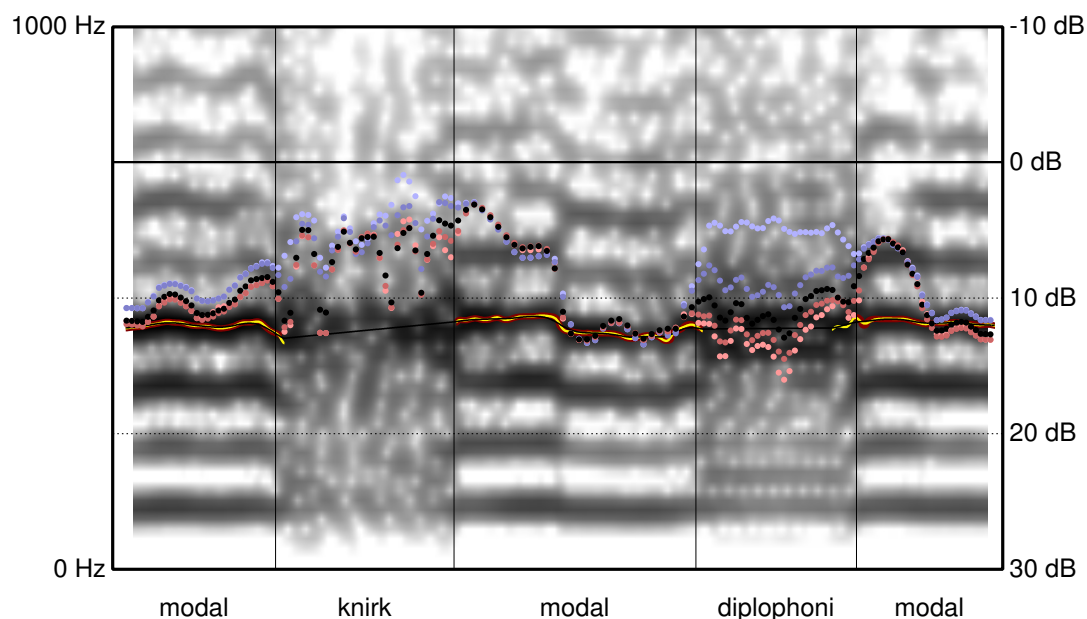
For optagelserne i serie I, II og III ses meget beskedne forskelle mellem de fem kurveforløb. Analysen af [a] med varierende stemmekvalitet er gengivet i figur 6.4 på næste side. Forskellene er så beskedne at de to rødlige og de to blålige kurveforløb i stort omfang dækker for hinanden. Kun i optagelsen fra serie IV, der indeholder diplofoni og knirk, optræder der væsentlige forskelle (se figur 6.5, næste side). Når der forekommer diplofoniske svingninger, resulterer bredere filterbånd (blålige) i højere-liggende H1:H2-kurver (dvs. mere energi i det øvre frekvensbånd), og modsat giver smallere filterbånd (rødlige) lidt lavereliggende kurver. I den passage hvor der er knirk, kan man også se nogle forskelle afhængig af bredden af de analyserede frekvensbånd, men det er af mindre betydning fordi H1:H2-målet ikke er meningsfuldt i de passager hvor signalet ikke er periodisk.

Der ser altså ud til at være relativt vide rammer for hvor brede eller smalle frekvensbånd der kan bruges i beregningen af H1:H2 uden at det fører til væsentligt forskellige H1:H2-værdier. Man ser dog også at det er uhensigtsmæssigt med alt for brede frekvensbånd hvis der er energi imellem deltonerne, fx på grund af diplofoni.

41 Jævnfør oversigten over de forskellige serier af testoptagelser på foregående side.



Figur 6.4. [a], varierende kompression fra underkomprimeret til overkomprimeret. De omtrentlige grænser mellem underkomprimeret, modal og overkomprimeret markeret med en lodret streg. H1:H2 beregnet ud fra frekvensbånd af varierende bredde; lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.



Figur 6.5. [æ] med tre typer stemtheds: modal, knirk og diplofoni. H1:H2 beregnet ud fra frekvensbånd af varierende bredde; lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.

6.15 Afprøvninger af H1:H2, BED og H1:H1½ robusthed ifht. f_0 -fejl

Eftersom beregningen af H1:H2, BED og H1:H1½ afhænger af f_0 , vil de være følsomme for fejl i f_0 -estimeringen. For at få et indtryk af hvor stor betydning evt. afvigelser i bestemmelsen af f_0 har, beregnes H1:H2, BED og H1:H1½ for en del af ovennævnte testoptagelser dels ud fra den estimerede f_0 , og dels ud fra estimater der forskydes op eller ned i forhold til den (formodede) korrekte f_0 -værdi. Resultaterne vedrørende H1:H1½ gennemgås i forbindelse med afsnit 6.19 (s. 113) der handler om diplofoni.

6.15.1 H1:H2

Beregningen af H1:H2 testes med off-set på +/- 1/16 og 1/8 af f_0 -værdien, svarende til henholdsvis ca. 6 % og 12.5 %. Sådant som f_0 plottes i analyserne, ville en forskydning på 1/8 svare til at f_0 -kurven ville have ligget midt mellem 3. og 4. deltone, eller midt mellem 4. og 5. deltone – dvs. en fejl der vil være let at se når overtonestrukturen træder klart frem.

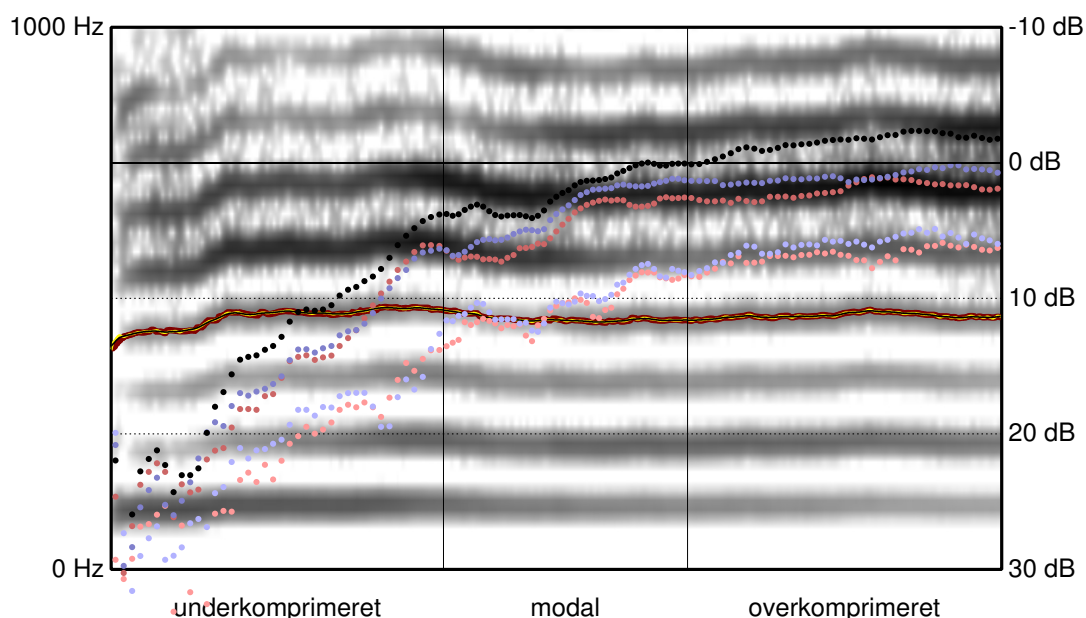
Farve	Offset		%
lys blålig	+ 1/8	+ 12.5	%
blålig	+1/16	+ 6.25	%
sort	f_0	0	%
rødlig	- 1/16	- 6.25	%
lys rødlig	- 1/8	- 12.5	%

Resultaterne findes i appendix D.I-IV. Ramme II for optagelsen med varierende kompression (vokalkvalitet [a], lydfil i appendix A1.2, resultat i appendix D.II) og optagelsen med knirk og diplofoni (lydfil i appendix A.IV, analyse i Appendix D.VI) er gengivet som figur 6.6 og 6.7. Som man kan se af figur 6.6 ligger det sorte kurveforløb der er baseret på de (formodenlig) korrekte f_0 -værdier, højest. Beregningen af H1:H2 på baggrund af et for højt eller lavt estimat af f_0 fører i begge tilfælde til en lavereliggende H1:H2-kurve. Dvs. der registreres relativt mere energi ved H1 når der ikke beregnes ud fra den "sande" f_0 . Det skyldes at H2 beregnes ud fra H1 (dvs. f_0) ved at gange med to. Det betyder at det frekvensbånd der tages i betragtning for at bestemme styrken af H2, forskydes dobbelt så langt væk fra den egentlige deltone som det frekvensbånd der tages i betragtning for at bestemme styrken af H1. Sammenhængen mellem fejlestimeringen af f_0 og H1:H2-niveauet er dog ikke monoton, og ved større forskydninger af f_0 -estimatet ville effekten vende. Ved mindre afvigelser er det dog overvejende sandsynligt at man vil se de højestliggende H1:H2-værdier (dvs. mest energi ved H2) når der beregnes ud fra den korrekte f_0 -værdi.

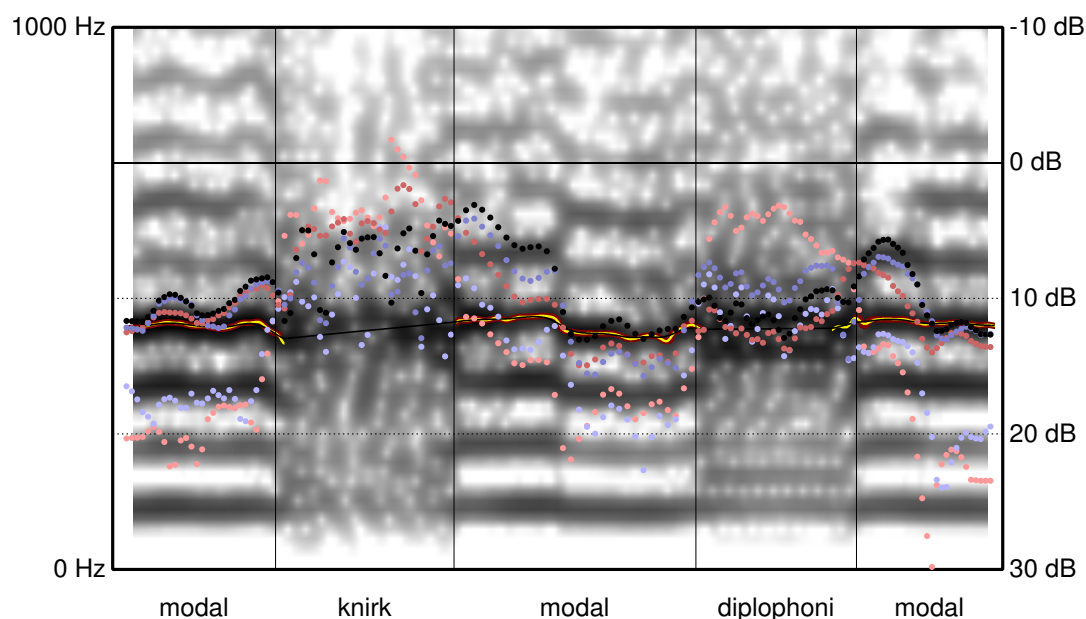
Forholdene bliver mere komplekse når signalet ikke er jævnt periodisk, se figur 6.7 næste side. I passagen med diplofoni fører en forskydning af f_0 -estimatet med -1/16 til en lavereliggende H1:H2 (mørk rødlig), mens en yderligere forskydning af f_0 -estimatet nedad til -1/8 fører til betydeligt højereliggende H1:H2-kurver (lys rødlig).

Det hænger sammen med at de sidebånd som modulationen tilføjer, gør at der nu ikke kun er energi ved de egentlige deltoner (H1, H2, H3 etc.), men også imellem dem. Dermed betyder en forskydning af f_0 ikke bare at det filterbånd der skulle fange fx H2, rammer ved siden af H2, men at det måske helt eller delvis rammer et af sidebåndene.

Som ventet ses det at H1:H2 er følsom for fejl i f_0 -bestemmelsen. Det ser ud til at en evt. *konstant* (relativt) fejlestimering af f_0 hovedsagelig ville medføre en forskydning af H1:H2-forløbene (nedad), mens kurveformen ville forblive omtrent uændret. I tilfælde hvor grundtoneestimatet svinger omkring den korrekte værdi, vil der dog ikke være tale om en konstant forskydning af kurveforløbet, og derfor vil der være grund til at være varsom med tolkningen af H1:H2-kurveforløbet. Ydermere ser det ud til at H1:H2-målet er mere følsomt for fejl-estimering af f_0 når der forekommer diplofoni. H1:H2-forløbet i diplofoniske passager må derfor tages med særlige forbehold for at det (sandsynligvis interpolerede) f_0 -estimat er tilstrækkeligt godt.



Figur 6.6. [a] med kompression varierende fra underkomprimeret til overkomprimeret. H1:H2 beregnet ud fra f_0 med offset af forskellig grad: lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %. Det fulde kurveblad findes som appendix D.II.



Figur 6.7. [æ] med tre typer stemthed: modal, knirk og diplofoni.

H1:H2 beregnet ud fra f_0 med offset af forskellig grad:

lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.

Det fulde kurveblad findes som appendix D.VI.

Imidlertid må man sige at generel over- eller underestimering af f_0 ikke er særlig sandsynlig; det er ikke sådan f_0 -algoritmer typisk fejler. Derimod er der netop risiko for fejl i bestemmelsen af f_0 i grænsetilfælde hvor signalet kun lige akkurat er tilstrækkelig periodisk til at f_0 -algoritmen leverer et bud på f_0 . Det kan fx ske i overgangen mellem modal og diplofonisk eller modal og knirket stemmekvalitet hvor der dermed er risiko for at der vil forekomme lokale fejl i estimeringen af f_0 . Der er derfor grund til at være særligt opmærksom på f_0 (og dermed tolkningen af H1:H2-forløbet) omkring passager med diplofoni eller knirk.

6.15.2 BED

Det viser sig at BED er så robust i forhold til fejl i f_0 sammenlignet med H1:H2 at det er nødvendigt med større offset af f_0 for at se hvornår BED overhovedet påvirkes af offset af f_0 . Følgende værdier benyttes:

Farve	Offset		%
lys blålig	+ 1/4	+ 25	%
blålig	+1/8	+ 12.5	%
sort	f_0	0	%
rødlig	- 1/8	- 12.5	%
lys rødlig	- 1/4	- 25	%

Ved et offset på 25 % kommer den beregnede grænse (på $1.5 \times f_0$) mellem det øvre og det nedre frekvensområde meget tæt på H1 eller H2:

$$f_0 \times 0.75 \text{ (offset)} \times 1.5 = 1.125 \times f_0$$

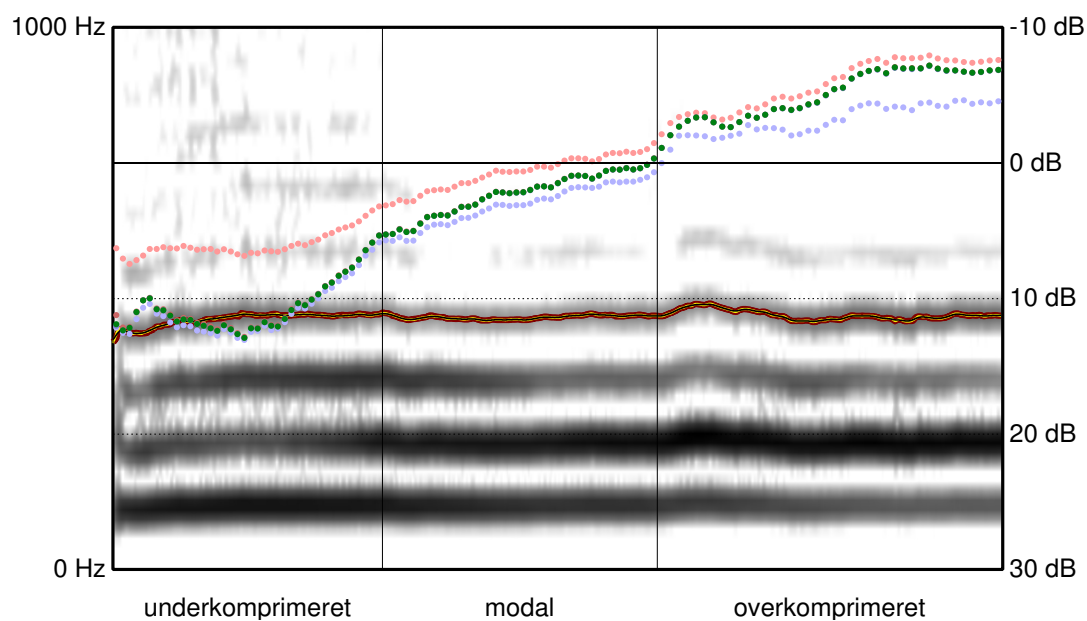
$$f_0 \times 1.25 \text{ (offset)} \times 1.5 = 1.875 \times f_0$$

Testresultaterne for BED findes i appendix E.I-E.VI. Ramme II for [i]-kompression (lydfil A.I.1) og optagelsen med knirk og diplofoni (lydfil A.IV.1) er gengivet som figur 6.8 og 6.9 på næste side. Generelt resulterer et offset på +25 % (lys blålig) i lidt lavereliggende BED-kurver, dvs. mere energi i det nedre frekvensbånd, mens et offset på -25 % (lys rødlig) resulterer i lidt højereliggende kurver, dvs. som udtryk for relativt mere energi i det øvre frekvensbånd.

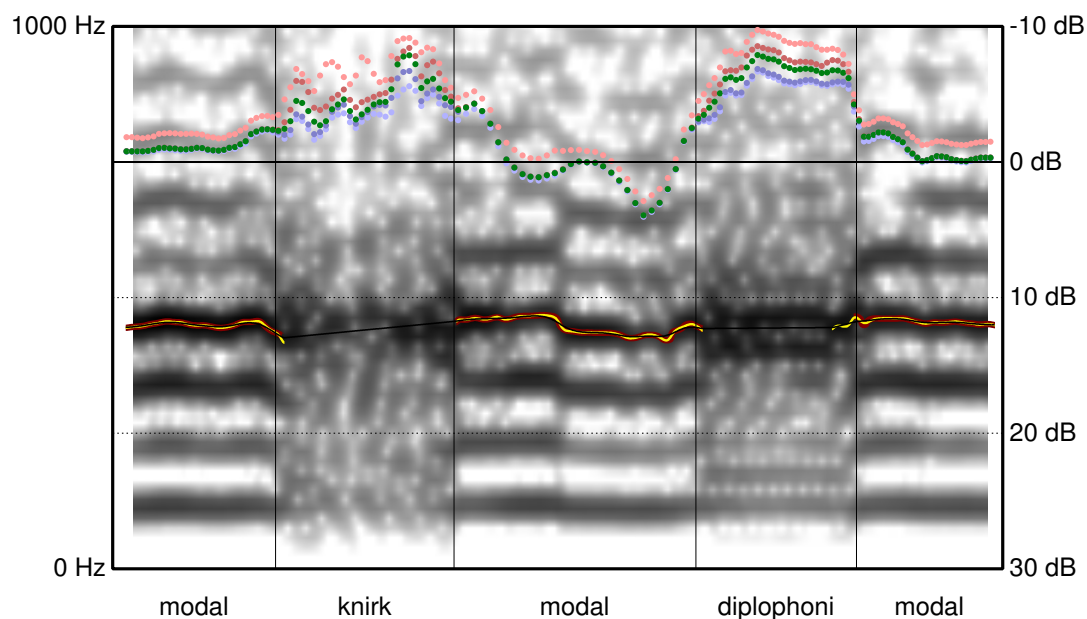
Der er dog nogle omstændigheder der ser ud til at føre til lidt større forskelle. For optagelsen af [i] med varierende kompression (figur 6.8) ses fx at en underestimering af f_0 på 25 % (lys rød) giver markant højere BED-kurve når stemmen er underkomprimeret, og en overestimering af f_0 på 25 % (lys blå) giver lidt lavere kurveforløb når stemmen er overkomprimeret.

Det er dog ikke videre sandsynligt at se f_0 -fejl i størrelsesordenen +/- 25 % med mindre der er uregelmæssigheder i periodiciteten, så jeg vil gå direkte til testen på optagelsen med knirk og diplofoni, se figur 6.9 næste side. I forbindelse med diplofoni eller knirk ses lidt større forskelle end i det omgivende materiale med modal stemthed. I passagen med knirk er der selvfølgelig ikke tale om en forskydning i forhold til en egentlig f_0 , men alene i forhold til den interpolerede værdi mellem de to omgivende periodiske passager.

Det er kun i forbindelse med diplofoni (og knirk) at der overhovedet ses eksempler på at en forskydning på 12.5 %, har en synlig effekt. Der er dog tale om en forholdsvis beskeden forskydning. BED viser sig således at være betydeligt mere robust i forhold til fejlestimering af f_0 sammenlignet med H1:H2, ikke mindst i forbindelse med diplofoni.



Figur 6.8. [i] med kompression varierende fra underkomprimeret til overkomprimeret. BED beregnet ud fra f_0 med offset af forskellig grad: lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %. Det fulde kurveblad findes som appendix E.I.



Figur 6.9. [æ] med tre typer stemthed: modal, knirk og diplofoni. BED beregnet ud fra f_0 med offset af forskellig grad: lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %. Det fulde kurveblad findes som appendix E.VI.

6.16 Test af overensstemmelse og afvigelser mellem H1:H2, BED og CoG

6.16.1 Serie 1, varierende stemmekvalitet, [i] og [a], middel grundtone.

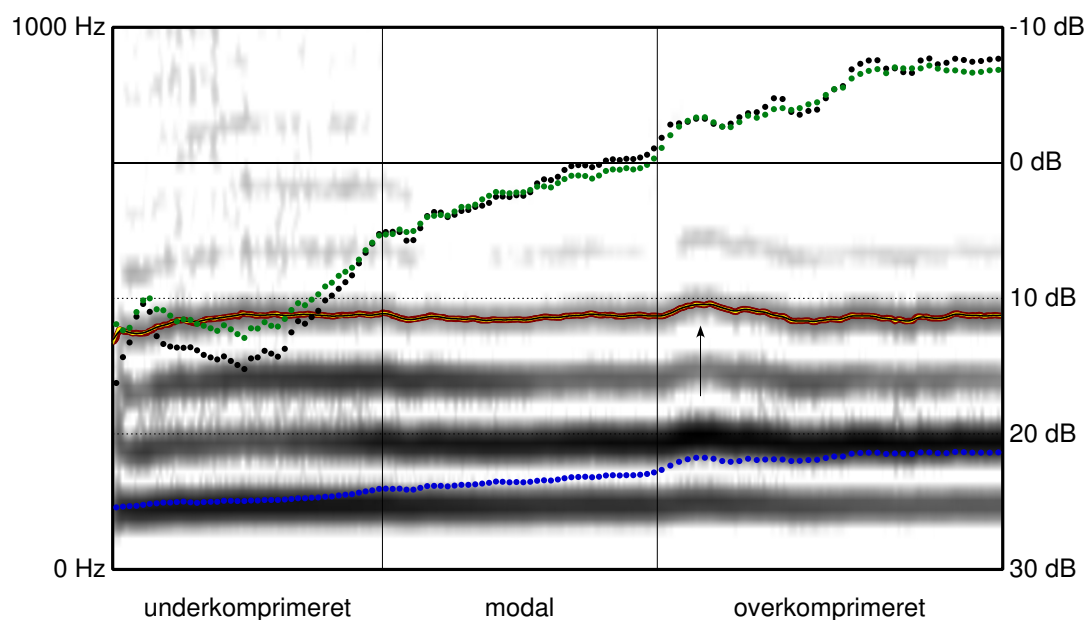
Første test går ud på at se om de tre mål der kan relateres til graden af kompression, dvs. H1:H2, BED, og COG, reagerer som ventet på ændringer i kompressionen når andre faktorer så vidt muligt holdes konstant. Vi starter derfor med at se på to testoptagelser hvor stemmekvaliteten varieres gradvist fra underkomprimeret (og let luftfyldt) over modal til overkomprimeret, mens grundtonen og vokalkvaliteten så vidt mulig holdes konstant. Figurerne 6.10 og 6.11 er baseret på testoptagelser med vokalkvaliteterne [i] og [a].

Overordnet set viser både H1:H2 (sort), BED (grøn) og CoG (blå) efter hensigten en opadgående tendens fra underkomprimeret til overkomprimeret. Hermed ses det at under i øvrigt konstante forhold kan H1:H2, BED og CoG alle bruges til at illustrere ændringer i stemmens kompression.

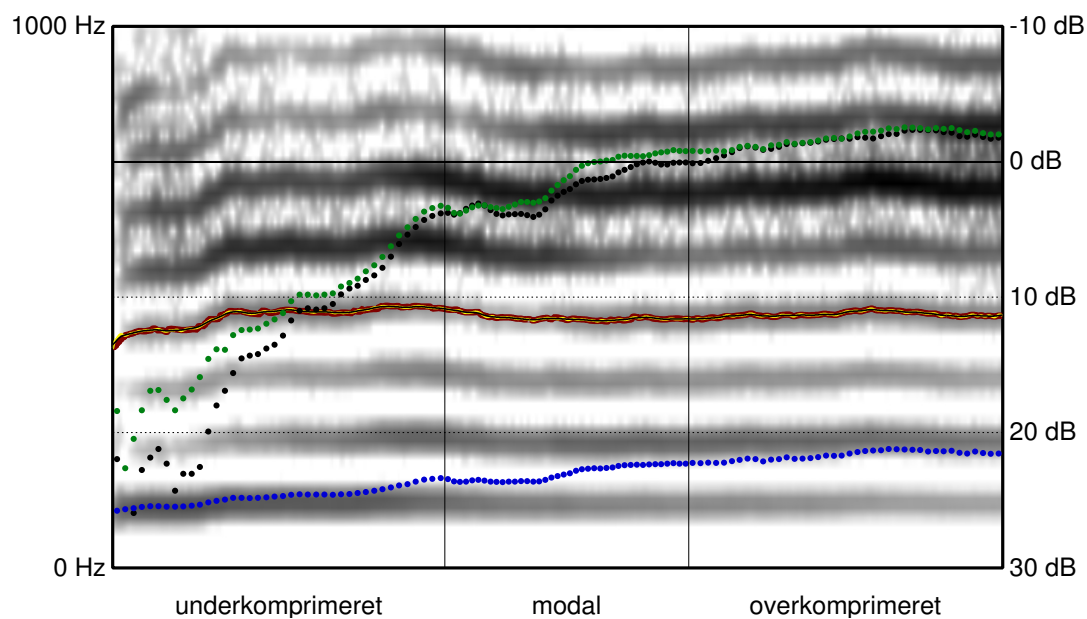
Selv om de tre mål samstemmende følger ændringen i kompression er stigningstakten ikke fordelt på samme måde for de tre mål. På figur 6.11 ser man fx at H1:H2 og BED stiger mere stejlt i omtrent den første tredjedel af forløbet (markeret som underkomprimeret) end i de sidste to tredjedele, hvorimod CoG tilsyneladende stiger mere jævnt igennem hele forløbet.

H1:H2-, BED- og CoG-forløbene er også præget af små lokale fluktuationer som tidsmæssigt falder sammen med fluktuationer i fx f_0 -kurven. Fluktuationerne afspejler at taleren (naturligvis) ikke har været i stand til at producere et fuldstændig jævnt forløb hvor det udelukkende er kompressionen der varierer.

I testoptagelsen med [i] (figur 6.10) kan man se tegn på at fluktuationer i grundtonen, kan påvirke de tre mål. For eksempel slår grundtonen et lille sving opad kort efter grænsen mellem modal og overkomprimeret (markeret med en sort pil) som man kan se afføder tilsvarende sving opad for både H1:H2, BED og COG. Det bekræfter at der som ventet er grund til at være mere opmærksom på udsving i grundtonen når F1 ligger lavt. Det er derfor oplagt som det næste at kigge på de testoptagelser hvor det er grundtonen som varieres.



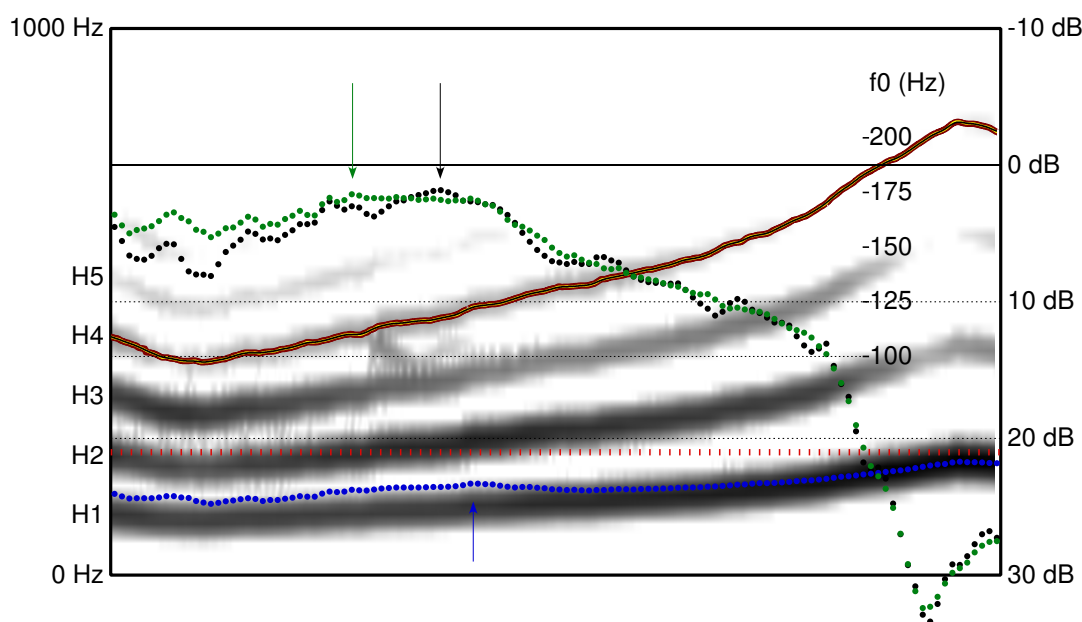
Figur 6.10. [i], varierende kompression fra underkomprimeret til overkomprimeret. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå). Det fulde kurveblad findes som appendix F.I.



Figur 6.11. [a], varierende kompression fra underkomprimeret til overkomprimeret. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå). Det fulde kurveblad findes som appendix F.II.

6.16.2 Serie 2, varierende grundtone, [i] og [a], middel kompression.

I de følgende to optagelser ser vi på [i] og [a] udtalt med en grundtone der varierer fra lav til høj, mens stemmeklang og vokalkvalitet så vidt muligt er holdt konstant. I almindelig tale vil grundtonen for mænd som en tommelfingerregel falde inden for området ca. 90-150 Hz, og det er derfor vigtigst hvad der sker i dette interval. (I forbindelse med denne undersøgelse kommer jeg kun til at se på mandlige talere).



Figur 6.12. [i], varierende grundtone. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå). Den anslåede F1-værdi er markeret med en stiple rød linje. Det fulde kurveblad findes som appendix F.III.

Det fremgår af 6.12 at H1:H2-målet (sort) varierer betydeligt afhængig af grundtonen. Grundtonen (H1) når den laveste værdi på omkring 90 Hz til venstre i forløbet, og topper ved ca. 150 Hz til højre på figuren. H2 bevæger sig følgelig fra ca. 180 Hz til ca. 300 Hz. Som det er karakteristisk for [i] ligger vokalens første formant (F1) relativt lavt ved ca. 225 Hz. Den omtrentlige placering af F1 er markeret i figuren med en tyk rød stiple linje⁴².

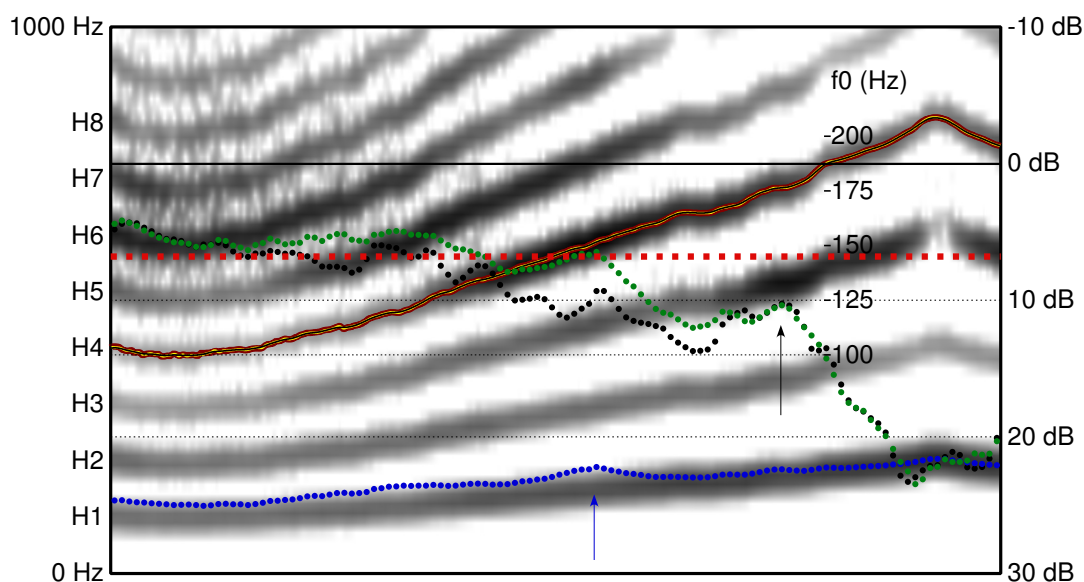
Overordnet udviser H1:H2 et stigende-faldende forløb som funktion af stigende f_0 . H1:H2-forløbet stiger efterhånden som grundtonen stiger fra sin mindste værdi og op til ca. 115 Hz hvor H1:H2-forløbet topper (markeret med en sort pil). Efterhånden som grundtonen stiger nærmer H2 sig til F1, og H1:H2-kurven topper derfor (omtrent) der hvor H2 frekvensmæssigt falder sammen med F1, hvorved H2 forstærkes mest i forhold til H1. Over ca. 115 Hz vender forløbet, og efterhånden som grundtonen bevæger sig højere op nærmer H1 sig til F1, mens H2 bevæger sig væk

⁴² Her vurderet på øjemål, og stiliseret angivet som en ret linje.

fra F1. Dermed sker der en forskydning i retning af energi ved H1 som resulterer i et nedadgående H1:H2-forløb.

BED-forløbet (grøn) udviser også et stigende-faldende forløb som funktion af stigende f_0 . BED-forløbet toppe ved en lidt lavere f_0 på ca. 100 Hz (markeret med en grøn pil), og ved frekvenser under ca. 125 Hz ser BED ud til ikke at blive helt så kraftigt påvirket af ændringer i f_0 som H1:H2. Når grundtonen er over ca. 125 Hz, ser H1:H2 og BED ud til at blive påvirket på nogenlunde samme måde af ændringer i grundtonen og begge falder kraftigt når f_0 når over ca. 160 Hz. Bemærk at når H3 er over 500 Hz, dvs. når f_0 er over 166 Hz, er der kun to deltoner i frekvensområdet 0-500 Hz, som er det frekvensområde der tages i betragtning ved beregningen af BED-værdierne. Det betyder at der reelt ikke er nogen forskel på BED og H1:H2 når f_0 er højere end 166 Hz.

CoG (blå) udviser overordnet set et svagt stigende forløb for stigende f_0 , men med en lokal bred top i området omkring ca. 125 Hz (markeret med en blå pil) der formodentlig skyldes at H2 forstærkes af F1.



Figur 6.13. [a], varierende grundtone. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå). Den omtrentlige F1-værdi er markeret med en stiple rød linje. Det fulde kurveblad findes som appendix F.IV.

Ved [a], se figur 6.13, ses overordnet et faldende forløb for H1:H2 og BED, men med lokale udsving. H1:H2 og BED falder kraftigt når f_0 når over ca. 175 Hz (markeret med en sort pil). CoG bevæger sig overordnet set opad med stigende f_0 , men har en markant top ved $f_0 = 150$ Hz (markeret med en blå pil) som måske har at gøre med at H3 her ligger så højt som den kan, før den begynder at bevæge sig uden for det frekvensområde som tages med i beregningen af CoG (som er 0-500 Hz). Da både

H1:H2 og BED også har en lille top på dette sted, er det dog ikke utænkeligt at det skyldes et lille udsving i stemmens kompression.

Sammenholder man resultaterne for de to vokalkvaliteter kan man se at H1:H2 og BED overordnet set påvirkes mere af ændringer i grundtonen når F1 ligger lavt. Det ser også ud til at H1:H2 er lidt mere følsom for ændringer i f_0 end BED er. Grundtonens indvirkning på CoG-værdien ser ud til at være omtrent den samme for de to vokalkvaliteter. Selvom H1:H2 og BEDs udsving som følge af ændringer i f_0 især inden for det primære grundtoneområde fra 90-150 Hz er mindre for [a] end de er for [i], så er udsvingene måske mere abrupte for [a]. Dermed kunne det se ud til at det ikke entydigt er en fordel med en højereliggende F1.

6.16.3 Serie 3, varierende vokalkvaliteter, middel grundtone, middel kompression.

I de følgende optagelser varieres vokalkvaliteten, mens grundtone og stemmeklang så vidt muligt holdes konstant. Lydoptagelserne findes som appendix A.III.1, A.III.2 og A.III.3.

I figur 6.14 ses H1:H2, BED og CoG for optagelse A.III.1 der indeholder en serie urundede fortungevokaler med stigende åbningsgrad fra det snævre [i] til det åbne [a]⁴³. I takt med åbningsgraden stiger F1 fra ca. 225 Hz ved [i] til omtrent 750 Hz ved [a]. Effekten af F1 kan ses i smalbåndsspektrogrammet ved at den eller de deltoner der forstærkes mest af F1, træder mørkere frem. Tilsammen tegner deltonerne billedet af det opadgående F1-bånd fra venstre mod højre.

H1:H2 gennemløber et stigende-faldende forløb efterhånden som F1 stiger, og når den højestliggende værdi (markeret med en sort dobbelpil) der hvor F1 krydser H2 og dermed forstærker H2 mest i forhold til H1. Den relativt beskedne ændring i vokalkvalitet (åbningsgrad) fra den snævre [i] til den lidt mere åbne [e] har ret dramatisk indvirkning på H1:H2-forholdet: Den lavest- og den højestliggende H1:H2-værdi ses i hhv. [i] og [e]! Det giver derfor endnu en indikation i retning af at snævre vokaler er mindre favorable i forhold til at se på H1:H2-forholdet. Efterhånden som F1 stiger højere op og dermed igen væk fra H2, falder H1:H2-kurven ned i et lidt lavere leje, og ændringen i åbningsgrad fra [æ] til [a] hvor F1 i dette tilfælde bevæger sig mellem H4 og H5, har ikke nogen nævneværdig indflydelse på H1:H2-forholdet.

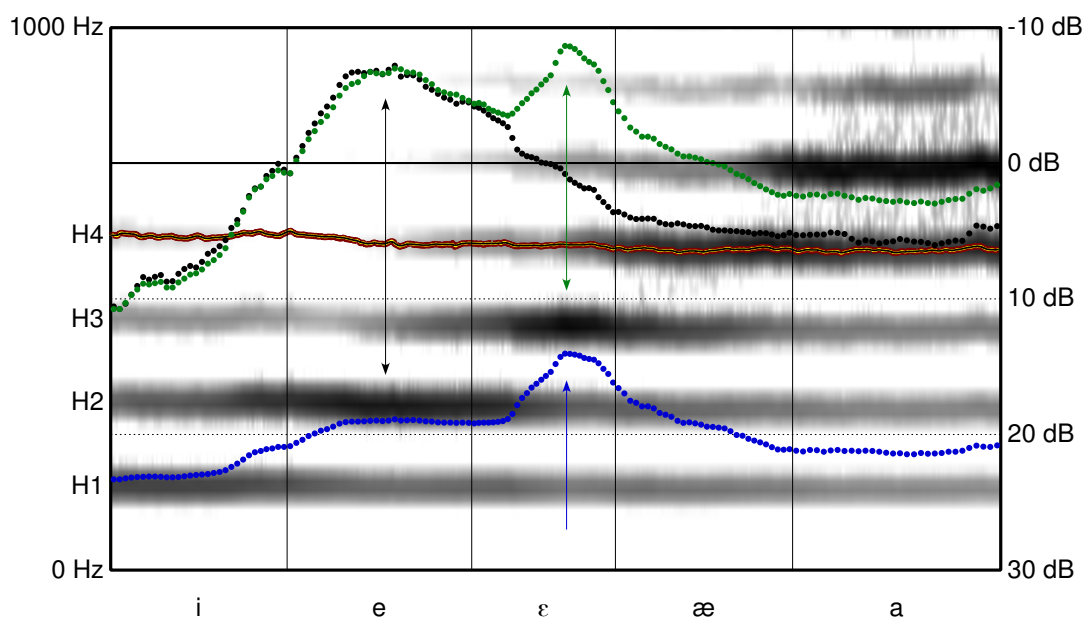
BED gennemløber et lidt mere kompleks forløb. Frem til overgangen mellem [e] og [ɛ] følges BED med H1:H2. Idet F1 bevæger sig fra et niveau omkring H2 (markeret med sort pil) til et niveau omkring H3 (markeret med grøn pil), falder BED først lidt og stiger derefter til et endnu højere niveau, der hvor F1 falder sammen med H3 (markeret med grøn pil). Efterhånden som F1 stiger højere op og dermed ud af det frekvensområde der tages i betragtning ved beregningen af BED (0-500 Hz), falder BED-kurven ned i et lidt lavere leje.

CoG følger et tilsvarende forløb som BED, og når også sit højeste niveau der hvor F1 falder sammen med H3 (markeret med blå pil). Skiftet fra [i] til [e] influerer også på CoG, men forskellen i CoGs niveau fra [i] til [e] er af samme størrelsesorden som den forskel der ses mellem [e] og [ɛ]. CoG er således lidt mindre følsom for ændringer i det vokalområde hvor H1:H2 og BED udviser den største følsomhed for ændringer i vokalkvalitet.

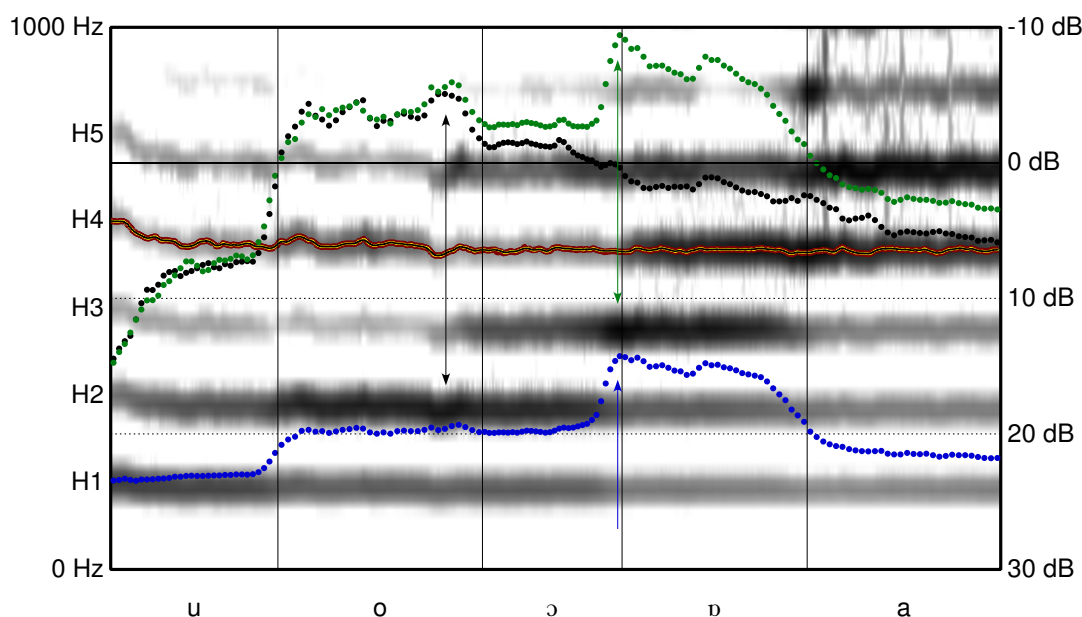
Figur 6.15 viser H1:H2, BED og CoG for optagelse A.III.2 der indeholder en serie rundede bagtungevokaler med stigende åbningsgrad fra det snævre [u] til det åbne [ɒ], plus det urundede men mere åbne [ɑ]. Som i foregående figur stiger F1 i takt med åbningsgraden fra ca. 225 Hz ved [u] til omtrent 750 Hz ved [ɑ].

43 NB! [ɑ] figurerer fejlagtigt som maskinskrift-a i figurene og de fulde kurveblade i appendix.

Bortset fra at ændringerne forekommer lidt mere abrupt i denne optagelse (fordi ændringerne i vokalkvalitet sker lidt mere abrupt), følger både H1:H2, BED og CoG samme forløb som for de urundede vokaler: H1:H2 når det højeste niveau der hvor F1 falder sammen med H2 som her er i vokalen [o]. Der er en lokal top i sidste del af [o], markeret med en sort pil, som falder sammen med et lille dyk i f_0 . Enten er det dykket i f_0 der afstedkommer toppen, eller også er der simpelthen et lille udsving i kompressionen på det sted. BED har også en top der hvor F1 falder sammen med H2, men når sit højeste niveau der hvor F1 passerer H3, ved overgangen fra [ɔ] til [ɒ] (markeret med grøn pil). Tilsvarende ligger CoG lavest i [u] hvor F1 er tættest på H1, og højest der hvor F1 er tættest på H3 (markeret med blå pil).



Figur 6.14. Serie af urundede forlængede vokaler plus [a]. H1:H2 (sort), BED (grøn) og CoG (blå). Det fulde kurveblad findes som appendix F.V.



Figur 6.15. Serie af bagtungevokaler plus [ɑ]. H1:H2 (sort), BED (grøn) og CoG (blå). Det fulde kurveblad findes som appendix F.VI.

Forskellen på formantforløbet i 6.14 og 6.15 er at F2 for de urundede vokaler ligger langt over 1000 Hz, mens den for de rundede bagtungevokaler ligger i området omkring 5-800 Hz. (F2 kan kun anes svagt på smalbåndsspektrogrammet. Se evt. formantanalysen i ramme III i den fulde figur, appendix F.VI).

H1:H2-forløbet toppe i begge optagelser der hvor F1 når en værdi der forstærker H2 mest. Ændringen i styrkeforholdet mellem H1 og H2 og tilsvarende for BED i vokalområdet [i] til [e] og [u] til [o] er meget markant. For H1:H2-forholdet er variationen kraftigere end den er for hele det øvrige spektrum af åbningsgrader, uanset om det er for- eller bagtungevokaler der er tale om.

BED-forløbet udviser to toppe: Den første er sammenfaldende med H1:H2-forløbets top der hvor H2 forstærkes af F1. Den anden top, som er kraftigere, forekommer der hvor F1 forstærker H3. Ved fortungerækken falder det sammen med [ɛ], ved bagtungerækken sker det ved overgangen mellem [ɔ] og [ɒ].

CoG stiger med stigende F1 så længe F1 er under 500 Hz. CoG toppe ved [ɛ] i fortungerækken og ved overgangen til [ɒ] i bagtungerækken.

I figur 6.16 som viser analysen af den sidste af de tre optagelser med varierende vokalkvalitet, veksles mellem [i], [y] og [u] – dvs. der er tale om snævre vokaler som er henholdsvis urundet og rundet, og henholdsvis fortunge og bagtunge. Da de alle er snævre vokaler, ligger F1 nogenlunde konstant lavt, mens F2 og F3 ændres som følge af skiftet fra urundet til rundet og skiftet fra for- til bagtunge; se formantforløbet i

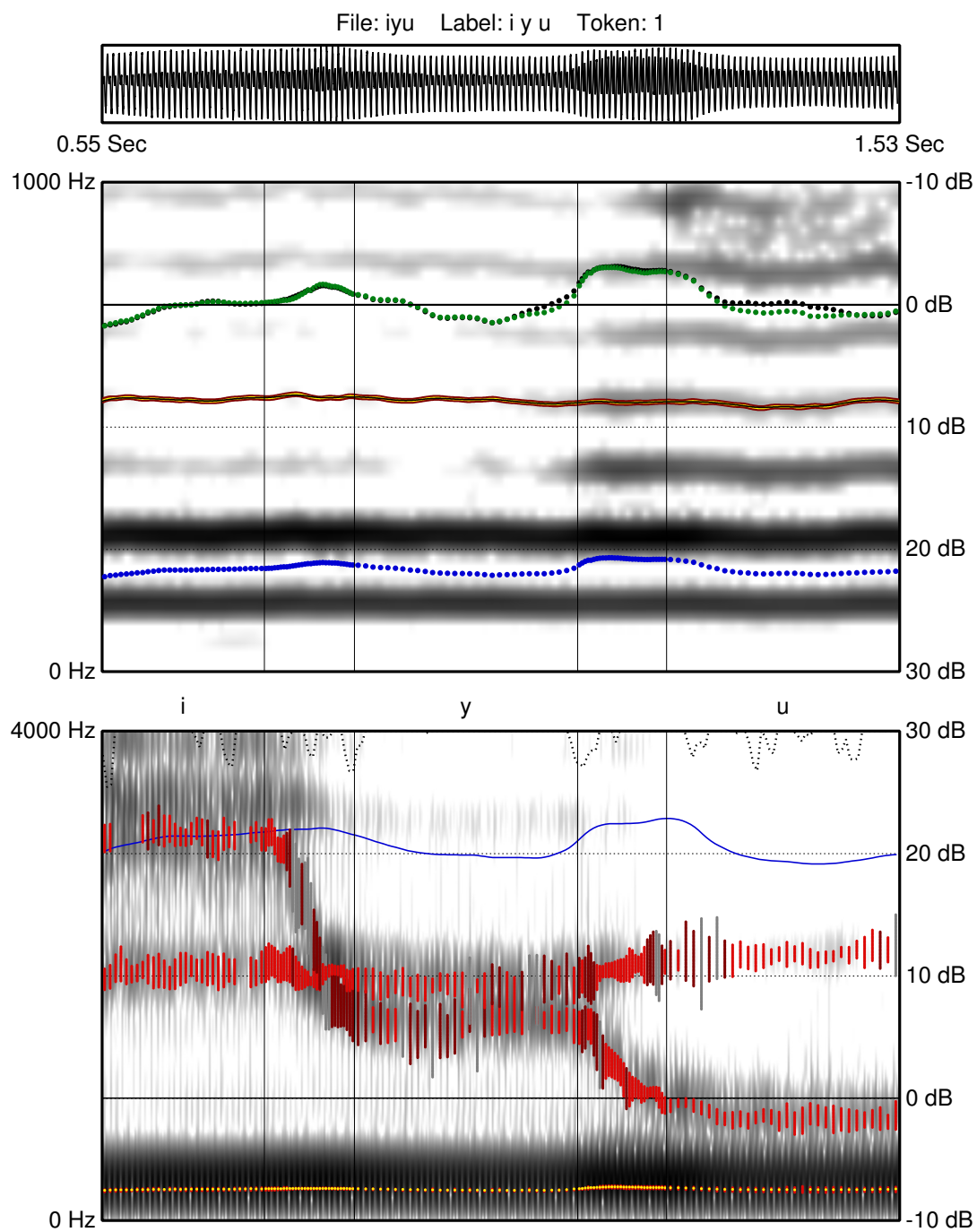
ramme III. F2 falder lidt fra [i] til [y] og noget mere til [u]. F3 falder fra [i] til [y] og stiger en anelse til [u]. Ændringerne i F2 og F3 har ikke nogen nævneværdig indflydelse på hverken H1:H2, BED eller CoG. De udsving i H1:H2, BED og CoG der ses fx i overgangen fra [i] til [y] og fra [y] til [u], skyldes nærmest usynlige udsving i F1! Det er ikke indlysende ved at se på figur 6.16 at det forholder sig sådan, men ændringerne i F1 kan anes i ramme III, tydeligst i overgangen mellem [y] og [u]. Sammenligninger af kurveforløbene fra flere tilsvarende optagelser bestyrker mig i den opfattelse at det er F1 som forårsager udsvingene. At så små udsving kan have konsekvenser for H1:H2, BED og CoG når F1 ligger lavt (i forhold til H1 og H2), giver endnu en indikation af at det vil være sikrere at få mål for kompressionen i mere åbne vokaler. Det giver også anledning til at overveje om F1 måske burde plottes i ramme II sådan at man lettere kunne se små ændringer, og samtidig hurtigt ville kunne se hvor F1 ligger i forhold til de nedre deltoner.

6.17 Opsummering af test

Det er indlysende ud fra den foregående test af grundtonens indflydelse på H1:H2, BED og CoG at i hvert fald H1:H2- og BED-forløbene i denne afprøvning af ændringer i vokalkvalitet ville have set anderledes ud hvis grundtonen havde ligget et andet sted. Selv om man derfor på ingen måde ville kunne danne en "korrektions-tabel" ud fra disse forsøg, så illustrerer de principperne bag de faktorer der spiller ind.

På baggrund af de gennemgåede testoptagelser må man konkludere, at for at H1:H2, BED og CoG-analyserne skal kunne bruges til at sammenligne stemmekvalitetsforholdene i to optagelser, er det vigtigt at både f_0 -værdierne og vokalkvaliteten i de to optagelser ikke adskiller sig væsentligt fra hinanden, og det må holdes for øje at der er kombinationer af vokalkvalitet og grundtonehøjde hvor selv meget små forskelle vil have en betydning.

Det er også tydeligt at der er en del tilfælde hvor de tre kurveforløb vil reagere forskelligt på de faktorer der vil kunne virke forstyrrende i forhold til at undersøge ændringer i stemmens kompression. Derfor vil samstemmende eller afvigende forløb for de tre nævnte mål kunne bruges til at vurdere hvor troværdige resultaterne er.



Figur 6.16. Serie af svævre vokaler: [i], [y] og [u]. H1:H2 (sort), BED (grøn) og CoG (blå).
Analyse af alternativ optagelse af samme serie af vokaler findes som appendix F.VII.

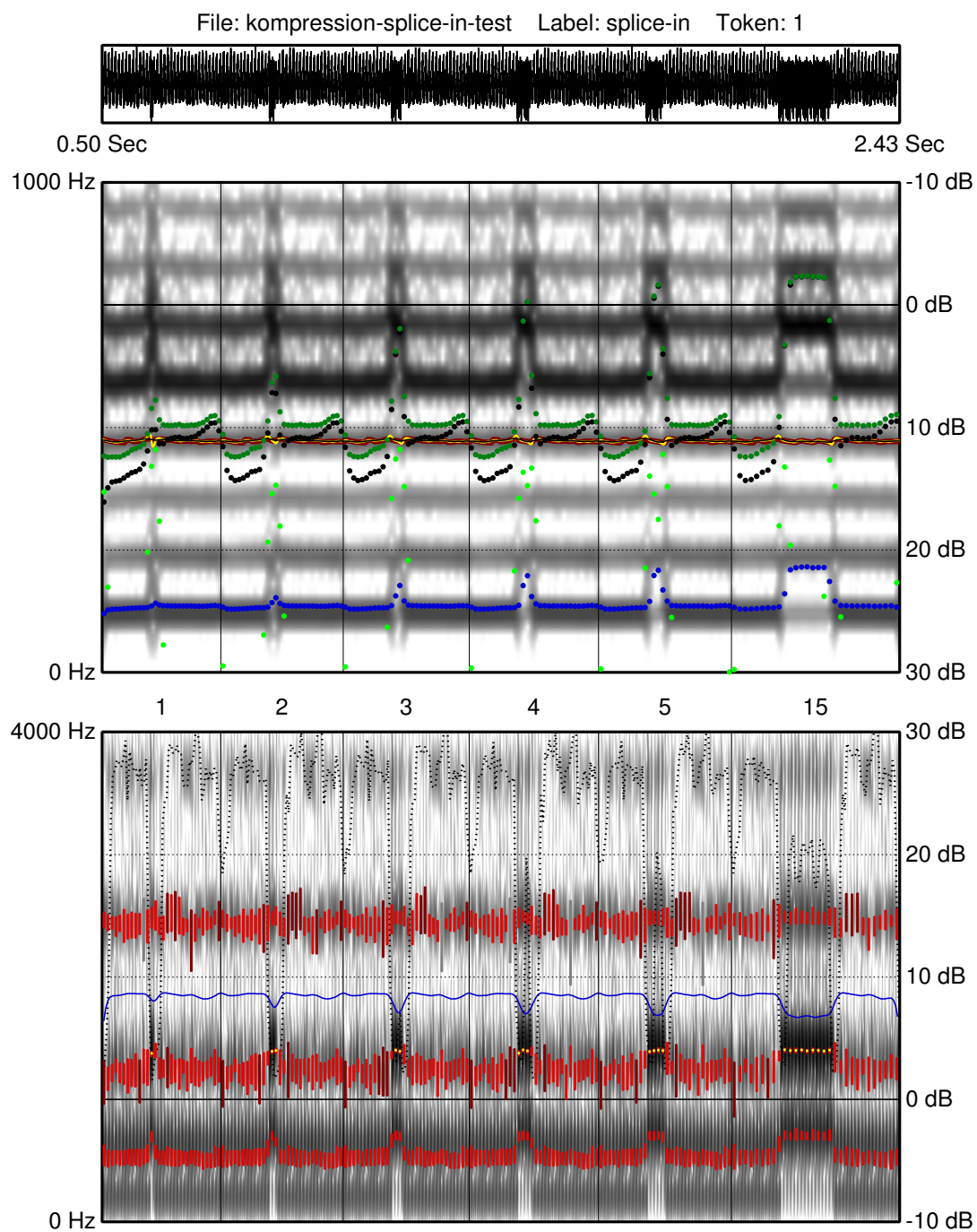
6.18 Er metoderne hurtige nok?

H1:H2 og BED beregnes på baggrund af en spektral analyse af signalet. Man kunne derfor forvente at de vil have samme tidsopløsning som den spektrale analyse selv. Bestemmelsen af H1:H2 og BED trækker dog foruden spektrogrammet også på et f_0 -estimat, og det kan ikke udelukkes at den måde målingen af H1:H2 og BED er implementeret på og den grafiske præsentation af målene, eventuelt kunne bidrage til at forringe tidsopløsningen. Det er derfor på sin plads at teste hvor hurtigt metoderne reagerer på relevante ændringer i stemmekvaliteten.

Jeg har konstrueret et testsignal på baggrund af den tidligere nævnte optagelse af [a]-vokal udtalt med gradvist stigende kompression som ligger bag figur 6.11, ved at klippe korte intervaller med høj kompression ud og sætte dem ind i et området med lav kompression. Jeg har udvalgt to intervaller med så vidt mulig samme grundtonefrekvens (specifikt ca. 117 Hz med et par Hz nøjagtighed); et interval på ca. 0.28 sekunder i begyndelsen af forløbet (med lav kompression⁴⁴) og et interval i den sidste del af forløbet (med høj kompression). I den komprimerede del har jeg segmenteret fem enkeltperioder ud sådan at der kan klippes sekvenser af 1-5 perioders varighed ud. I den ukomprimerede del har jeg identificeret et passende sted hvor de korte sekvenser kan indsættes. I denne procedure er der taget højde for at alle klippene så vidt muligt laves på samme sted i forhold til de glottale pulser, og at de er placeret ved en nulgennemgang i lydsignalet. Jeg har lavet 6 sekvenser med indskud af 1, 2, 3, 4, 5 og 15 perioders varighed, svarende til hhv. ca. 8.5, 17, 25, 34, 43 og 128 ms varighed.

Det fremgår af figur 6.17 at der kræves 4-5 svingninger før kurveforløbene når helt op på den target-værdi som man kan se i det længste interval og ned igen. Iagttagelsen gælder både H1:H2, BED og CoG. Det betyder at analysernes tidsopløsning giver mulighed for at følge forandringer som sker i løbet af ca. 20 ms, svarende til to til tre svingningsperioder. (Halvdelen af varigheden af 4-5 svingninger). For at sætte tallet i perspektiv kan nævnes at lange vokaler typisk vil have en varighed på mellem 60 og 200 ms. Tidsopløsningen lever lige akkurat op til det mål jeg fremsatte i starten af kapitel 6. Sammenligner man med smalbåndsspektrogrammet, ser det ud til at H1:H2, CoG og BED har sammenlignelig, hvis ikke samme, tidsopløsning.

⁴⁴ De spektrale karakteristika ændrer sig løbende og relativt hurtigt i den underkomprimerede del af den oprindelig optagelse. Selvom det er en relativt kort sekvens jeg har brugt til at sætte de komprimerede bidder ind i, så er de spektrale mål ikke konstante igennem denne sekvens, se figur 6.17.



Figur 6.17. Afprøvning af metodernes reaktionstid på baggrund af sammenklipping af sekvenser med høj og lav kompression. Længden af de mere komprimerede sekvenser øges gradvist fra 1 til 15 perioder.

6.19 Test af $H1:H1\frac{1}{2}$ – Afsøgning af diplofoni

Til at afprøve $H1:H1\frac{1}{2}$ benyttes testoptagelse af vokallyd med passager med diplofoni og knirk, se figur 6.18. I de modale passager træder deltonerne tydeligt frem. Jeg har markeret de nederste fire deltoner $H1$ til $H4$ til venstre for spektrogrammet. Diplofoni fremtræder i smalbåndsspektrogrammet som sidebånd, dvs. frekvensbånd mellem de egentlige harmoniske deltoner. Som helhed varierer styrken (sværtningen) af sidebåndene med graden af diplofoni. Jeg har markeret sidebåndet mellem $H1$ og $H2$ som $H1\frac{1}{2}$, mellem $H2$ og $H3$ som $H2\frac{1}{2}$, og mellem $H3$ og $H4$ som $H3\frac{1}{2}$. Det er ikke alle deltonerne og sidebåndene der træder lige tydeligt frem i dette eksempel på diplofoni. Fx er $H2$, $H2\frac{1}{2}$ og $H3$ noget udløst, og man kan kun se antydningen af det sidebånd der ligger under $H1$, ved modulationsfrekvensen, dvs. $\frac{1}{2} \times f_0$, markeret med $H\frac{1}{2}$. Deltonerne $H1$ og $H4$ og sidebåndene $H1\frac{1}{2}$, $H3\frac{1}{2}$ træder dog tydeligt nok frem til at der ikke er tvivl om at stemmelæbesvingningerne er periodiske, og at der er tale om diplofoni. Til sammenligning ser man ikke nogen deltoner i den knirkede passage hvor stemmelæberne altså ikke har fulgt et periodisk svingningsmønster.

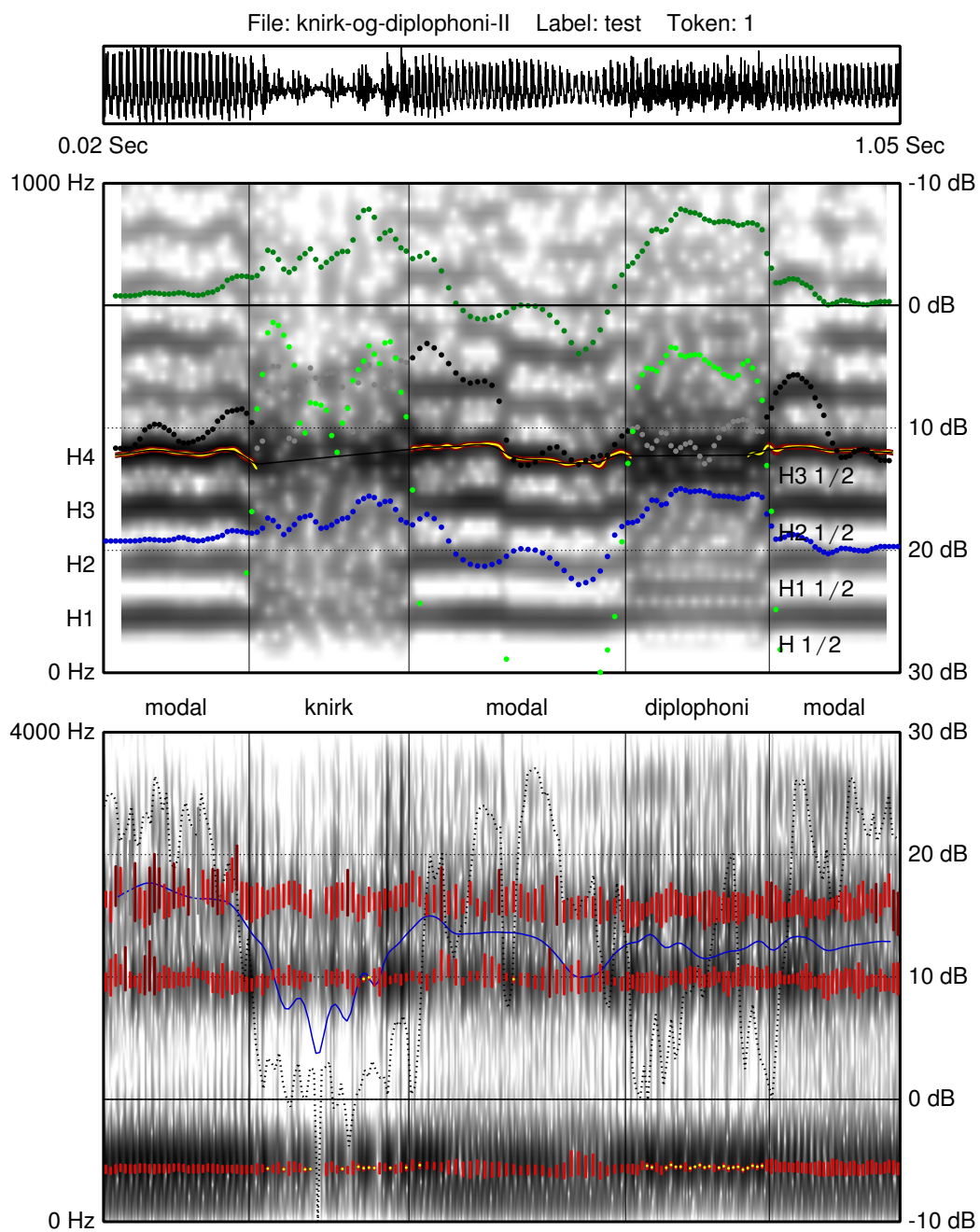
$H1:H1\frac{1}{2}$ -forløbet som plottes med den lidt lysere grønne, ligger stort set konstant så lavt i de modale passager at det falder uden for det område der plottes. Dvs. at styrken af $H1$ er mere end 30 dB kraftigere end energiniveauet ved $H1\frac{1}{2}$ (midt mellem $H1$ og $H2$). $H1:H1\frac{1}{2}$ stiger meget abrupt ved overgangen til passagen med diplofoniske svingninger, og ved toppen af kurveforløbet registreres energi ved $H1\frac{1}{2}$ som kun ligger 4-5 dB under styrken af $H1$. Testen viser altså at $H1:H1\frac{1}{2}$ -målet reagerer formålstjenligt på forskellen mellem modal og diplofonisk stemthed.

$H1:H1\frac{1}{2}$ giver dog ikke kun udsving i forbindelse med diplofoni. Kurveforløbet ligger også højt i den knirkede passage hvor det dog må pointeres at der egentlig ikke er tale om en måling af $H1$ og $H1\frac{1}{2}$ (fordi der hverken er deltoner eller sidebånd). Ud over at man kan se forskel på diplofoni og knirk i smalbåndsspektrogrammet, så er der også andre forskelle mellem den diplofoniske og den knirkede passage som det er værd at fremhæve: $H1:H1\frac{1}{2}$ fluktuerer mere i den knirkede passage. Harmonicity to Noise Ratio (HNR – sort stiplede linje i ramme III) ligger lavere i den knirkede passage som udtryk for at den diplofoniske passage registreres som mere periodisk end den knirkede passage. Der ses et dyk i intensiteten (blå linje i ramme III) i passagen med knirk hvorimod intensiteten ikke ændrer sig væsentligt i den diplofoniske passage sammenlignet med de modale passager.

Bemærk at CoG og BED (henholdsvis blå og grøn) ligger højere i både den knirkede og den diplofoniske passage, mens $H1:H2$ (sort, og grå hvor f_0 er interpoleret) ikke stiger tilsvarende i den diplofoniske passage. Jeg vil vende tilbage til det forhold i forbindelse med afsnit 6.20.

Nøjagtig som det gælder for $H1:H2$ -målet, er bestemmelsen af $H1:H1\frac{1}{2}$ også afhængig af et rimeligt bud på f_0 . Jo mere uregelmæssigt eller diplofonisk signalet er, jo mindre sandsynligt er det at pitch-tracking-algoritmen vil give et brugbart bud på f_0 .

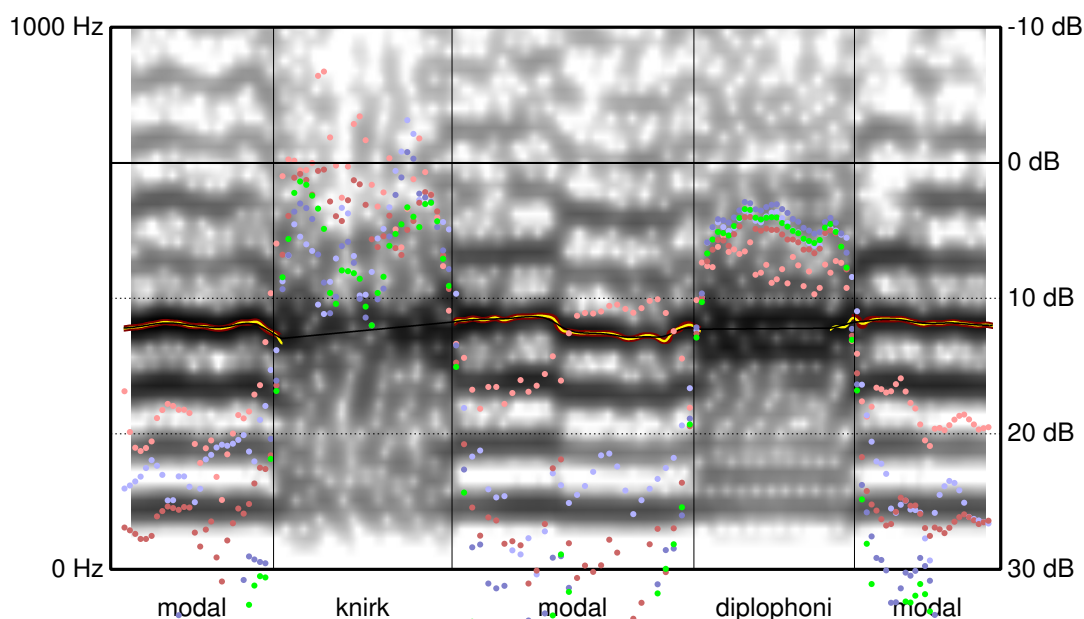
Der vil derfor højst sandsynlig forekomme "huller" i f_0 -trackingen i forbindelse med knirk og diplofoni. En måde at forsøge at overkomme den barriere på er ved at interpolere hen over de passager hvor f_0 ikke kan bestemmes. I outputtet kan de passager hvor f_0 er interpoleret, kendes ved at de brune og gule f_0 -kurver er fraværende, mens den sorte f_0 -kurve som viser den udglattede og interpolerede bud på f_0 står alene. I disse passager plottes H1:H2-forløbes samtidig med grå prikker i stedet for sorte.



Figur 6.18. Afprøvning af H1:H1½-målet på optagelse af [æ] med tre typer stemthed: modal, knirk og diplofoni.

Det er klart at jo længere stræk der interpoleres over, og jo mere bevægelse der er i f_0 , jo større risiko er der for at de interpolerede værdier ikke er brugbare. Det er derfor nødvendigt løbende at forholde sig til om de interpolerede f_0 -værdier er rimelige. For at få et indtryk af hvor stor betydning evt. afvigelser i bestemmelsen af f_0 har, testes H1:H1½ med off-set på +/- 1/16 og 1/8 af f_0 -værdien, svarende til henholdsvis ca. 6 % og 12.5 % efter samme princip som blev brugt for H1:H2 og BED i afsnit 6.15.

I figur 6.19 nedenfor gengives ramme II for afprøvningen på optagelsen med knirk og diplofoni som illustrerer effekten.



Figur 6.19. [æ] med tre typer stemthed: modal, knirk og diplofoni.

H1:H1½ beregnet ud fra f_0 med offset af forskellig grad:

lys blålig: + 12.5 %, blålig: + 6.25 %, grøn: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.

I en optagelse som denne – hvor der er tilstræbt så lidt variation i f_0 som muligt – er de interpolerede værdier formentlig så gode som de kan forventes at blive. I dette tilfælde ligger den interpolerede f_0 dog efter mit øjemål nok omkring 4-5 % for lavt: Den interpolerede f_0 ligger lavere end H4, men ikke så lavt at den falder midt mellem H4 og H3½ hvilket ville svare til at den lå 6.25 % for lavt. Det fører mig frem til at skyde på at fejlen er i omegnen af 4-5 %.

I den diplofoniske passage ses den største effekt ved en underestimering på 12.5 % (lys rødlig). I de modale passager er effekten af offsettet at H1:H1½-forløbet hæves så meget at det når inden for det område der plottes. Effekten af offsettet ser altså ud til at være større i de modale passager end i passagen med diplofoni, hvilket er heldigt: Det er sandsynligt at den interpolerede f_0 vil afvige noget fra den korrekte værdi, men

$H1:H1\frac{1}{2}$ lader til at være forholdsvis robust hvis afvigelsen fra f_0^{45} holder sig under ca. 12 % (en forskel der vil være relativt let at se).

6.20 Diplofoni: amplitude eller frekvensmodulation

Selvom foregående test illustrerede at $H1:H1\frac{1}{2}$ -målet ser ud til at reagere efter hensigten når man kan se tegn på diplofoni i smalbåndsspektrogrammet, så ville det være nyttigt hvis man kunne få en indikation af om den diplofoniske variation kun rammer amplituden af de enkelte pulser, eller om periodetiden også svinger mellem to forskellige værdier. Med andre ord: Hvis man kan afgøre om der er tale om amplitude- eller frekvensmodulation. Som nævnt vil $H1:H2$ blive påvirket hvis diplofonien også vedrører periodetiden – dvs. at der forekommer *frekvensmodulation* (med $f_0/2$) – og i det tilfælde vil det føre til en fejltagtig bedømmelse af stemmekildens kompression.

Jeg har eksperimenteret med at bruge Praats pulses-funktion til at undersøge variationen i timingen af de glottale pulser i forbindelse med diplofoni. Pulses-funktionen danner udgangspunkt for beregningen af jitter og shimmer, som er mål for henholdsvis ustabilitet i amplituden og timingen af de glottale pulser. Det viser sig at det er meget vanskeligt at undgå at funktionen springer hver anden puls over når der forekommer diplofoni, og i det hele taget lader funktionen til at fungere bedst med signaler der har en høj grad af periodicitet. Det betyder at pulses-funktionen ikke kan bruges til at afklare om det er amplituden eller periodetiden der varierer i diplofoniske passager.

Jeg vil derfor foreslå en anden fremgangsmåde: Ved at syntetisere vokaler med diplofonisk variation af enten amplituden eller periodetiden, altså amplitude- eller frekvensmodulation, kan man vise at de to typer modulation har distinkt forskellige konsekvenser for spektret. Det betyder at flere af undersøgelsens mål vil reagere forskelligt på de to typer modulation.

Syntetiseringen af eksemplerne er foretaget med et Praat-script skrevet af Niels Reinhold Petersen.⁴⁶ Scriptet giver mulighed for at generere vokaler med forskellige former for uregelmæssigheder i kildens svingningsmønstre. Jeg har lavet en mindre modifikation af scriptet så det kan producere et output der varierer gradvist fra en setting til en anden.

I det første eksempel (figur 6.20) introduceres diplofoni som *amplitudemodulation* med den halve frekvens af grundtonen: I det markerede interval dæmpes amplituden af hver anden periode gradvist. I begyndelsen af intervallet har alle pulser fuld amplitude. I slutningen af intervallet er amplituden af hver anden periode dæmpet helt. Det er kun amplituden som ændres; pulsformen holdes konstant. På smalbåndsspektro-

⁴⁵ Se i øvrigt afsnit 2.4.1 for en diskussion af begrebet grundtone i forbindelse med diplofoni.

⁴⁶ Scriptet var tidligere tilgængeligt på Niels Reinhold Petersens hjemmeside.

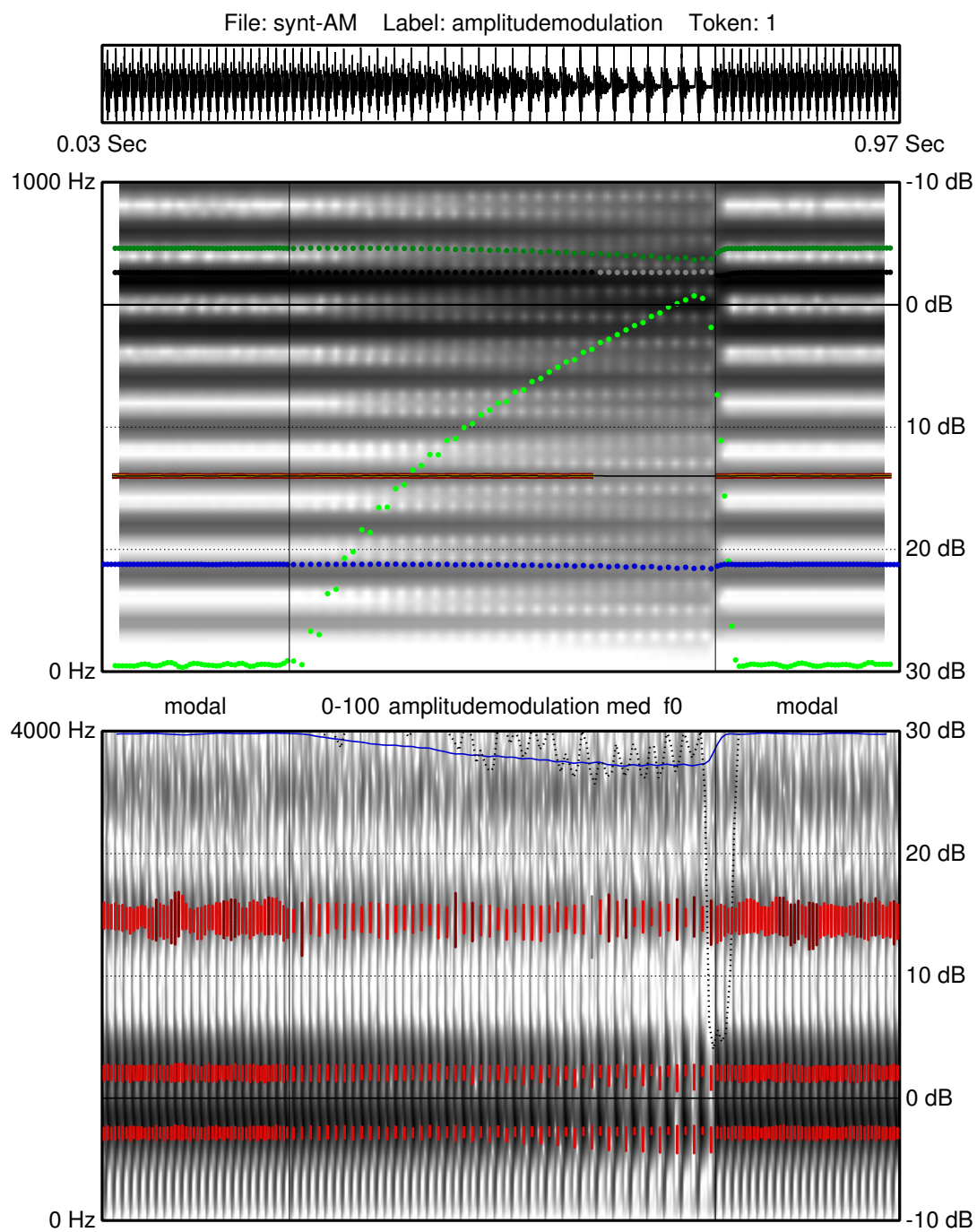
grammet kan man se at der gradvist optræder mere energi i nye frekvensbånd (sidebånd) mellem de oprindelige harmoniske deltoner.

Efter hensigten stiger $H1:H1\frac{1}{2}$ -kurven jo kraftigere modulationen bliver.

HNR ligger væsentlig højere i de umodulerede passager end man normalt vil se i optagelser af reel tale, og falder relativt lidt i forbindelse med amplitudemodulationen. For så vidt er det bemærkelsesværdigt at amplitudemodulationen får HNR til at falde. Amplitudemodulationen gør principielt ikke det syntetiserede signal mindre periodisk; set med matematiske briller får signalet en anden grundtone som er lig med modulationsfrekvensen.

$H1:H2$ påvirkes ikke af amplitudemodulationen. Det er i fuld overensstemmelse med beskrivelsen af de spektrale konsekvenser af amplitudemodulation der blev ridset op i afsnit 2.3.1. De sidebånd der genereres, falder udelukkende midt mellem de egentlige deltoner som altså ikke påvirkes.

CoG og BED forskydes en anelse i retning af energi i den nedre del af frekvensområdet. Det må tilskrives tilstedeværelsen af sidebåndet under $H1$ (som kunne betegnes $H1\frac{1}{2}$).



Figur 6.20. Syntetisk genereret vokal [a] med gradvist stigende amplitudemodulation.

I det andet eksempel (figur 6.21) introduceres *frekvensmodulation* med den halve frekvens af grundtonen. I det markerede interval forkortes henholdsvis forlænges periodetiden for hver anden periode gradvist. I begyndelsen af intervallet har de samme længde, i slutningen af intervallet er de korte perioder 40 % af den oprindelige periodelængde, de lange perioder forlænges så den samlede periodetid for to perioder holdes konstant.

Som ventet og som ved amplitude-modulation stiger $H1:H1\frac{1}{2}$ -kurven jo kraftigere diplofonien bliver. For de øvrige mål har frekvensmodulationen andre konsekvenser. Ser man på smalbåndsspektrogrammet, er det tydeligt at de sidebånd som frekvensmodulationen giver anledning til, ikke kun falder mellem de egentlige deltoner, men også falder sammen med disse og tilsyneladende giver anledning til både konstruktiv og destruktiv interferens. Resultatet er at der afhængig af modulationsstyrken ud over sidebånd mellem deltonerne opstår huller i deltonerækken, og at sidebåndene kan optræde væsentligt kraftigere end de nærmeste deltoner (se fx $H8$ og sidebåndet mellem $H7$ og $H8$ i figur 6.21).

HNR falder meget hurtigt til et niveau der ligger ca. 20 dB under hvad vi så ved det amplitudemodulerede signal. Jeg formoder at det lavere HNR-mål er et resultat af at der er "uorden" i overtonerækken pga. de huller der opstår, snarere end det skyldes signal uden for rækken af deltoner og sidebånd.

I modsætning til hvad vi så i det amplitudemodulerede eksempel, påvirkes $H1:H2$ af frekvensmodulationen, men ikke på nogen enkel måde: Efterhånden som modulationsstyrken øges, falder $H1:H2$ -kurven indtil den når et minimum der hvor man kan se et "hul" i $H2$ i smalbåndsspektrogrammet (ramme II). Derfra stiger $H1:H2$ -kurven indtil den ved den højeste modulationsstyrke når over det niveau som ses i den umodulerede del af signalet. Effekten ses også i smalbåndsspektrogrammet hvor $H1$ er meget svag, mens $H2$ har nået omtrent samme styrke som i det umodulerede signal.

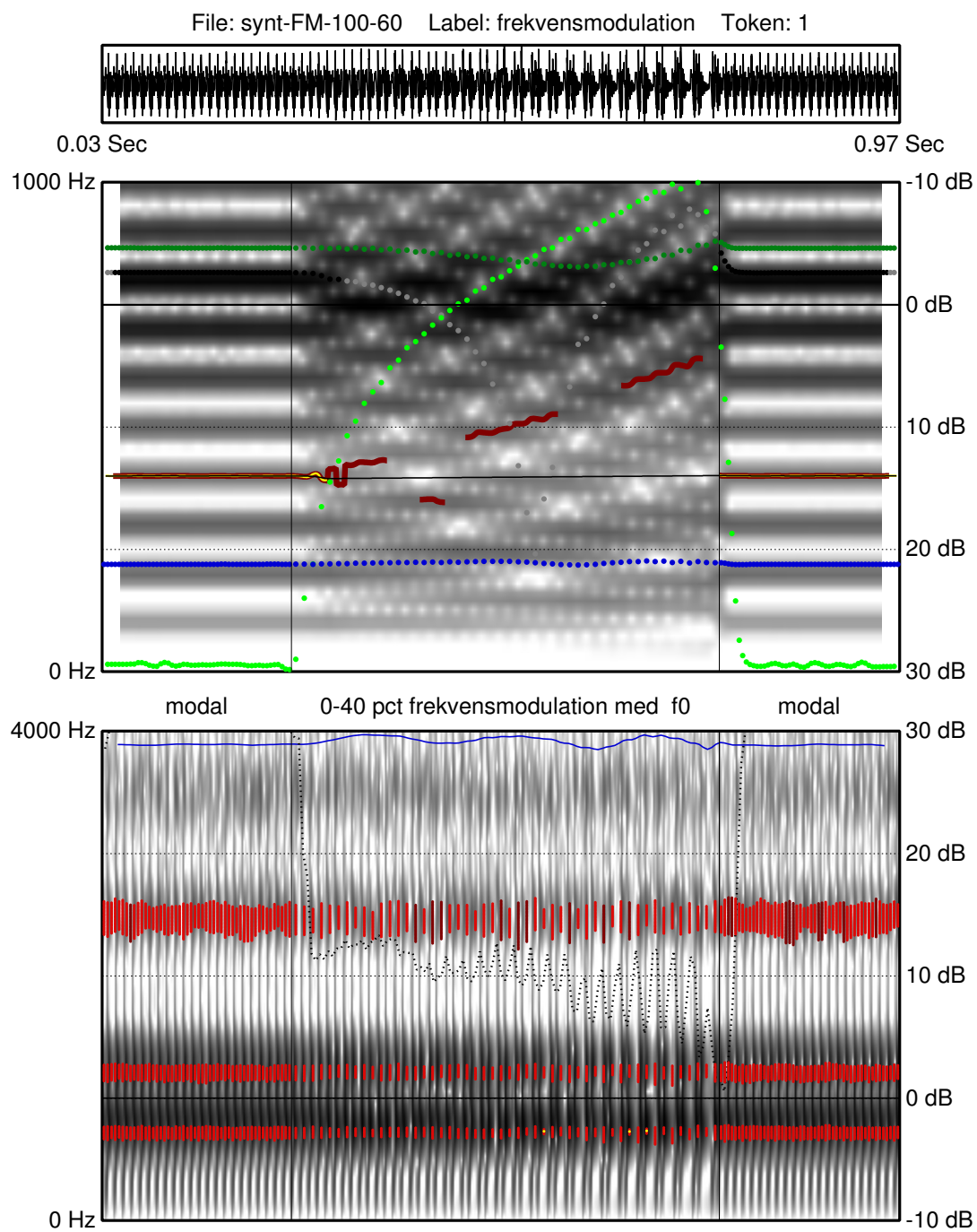
BED og CoG påvirkes også af frekvensmodulationen, men langt fra lige så voldsomt som $H1:H2$. CoG ser ud til at være det mål der påvirkes mindst.

I det sidste eksempel (figur 6.22) er der en kombination af amplitude og frekvensmodulation. Den maksimale styrke af begge er dog reduceret til hhv. 50 % reduktion af amplituden af hveranden puls og hhv. 25 % reduktion af periodetiden af hveranden periode (med en komplementær forlængelse af den efterfølgende periode). Essensen er at frekvensmodulationens effekt slår igennem og influerer på HNR, $H1:H2$, og i mindre omfang BED og CoG som om der kun havde været frekvensmodulation.

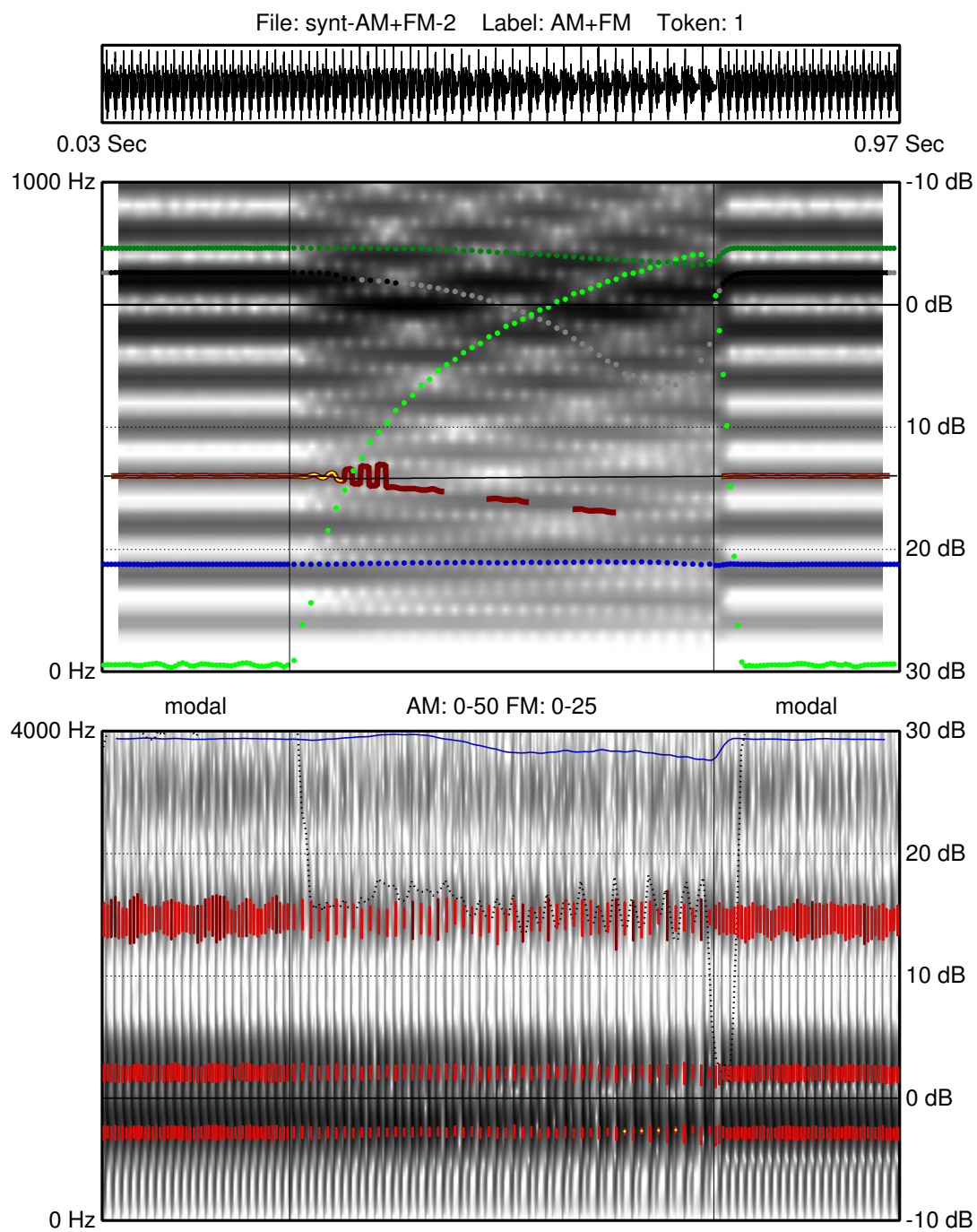
Det betyder at der er en række tegn som kan bruges til at vurdere om evt. diplofoni indebærer modulation af periodetiden: Hvis $H1:H2$ ser ud til at bevæge sig markant anderledes end BED og CoG, og hvis der ser ud til at forekomme huller i deltonerækken, så tyder det på frekvensmodulation. Det betyder at den diplofoni som vi så i figur 6.18, formodentlig indebærer modulation af periodetiden: Både $H2$ og $H3$ ser ud

til ret abrupt at blive svagere i den diplofoniske passage, og som konsekvens viser H1:H2 også et markant anderledes forløb end BED og CoG der begge udviser et stigende-faldende forløb henover den diplofoniske passage hvor H1:H2 ligger højere både umiddelbart før og efter diplofonien.

Forsøgene viser at amplitudemodulation ikke påvirker de tre mål, hvorimod frekvensmodulation påvirker H1:H2, men ser ud til at påvirke BED og især CoG mindre. Det betyder at CoG og til dels BED skulle kunne give en indikation af ændringer i kompressionen også i passager med uregelmæssige svingninger, uanset om det er amplituden og/eller periodetiden der fluktuerer.



Figur 6.21. Syntetisk genereret vokal [a] med gradvist stigende frekvensmodulation.



Figur 6.22. Syntetisk genereret vokal [a] med gradvist stigende amplitude- og frekvensmodulation.

7 Udformning af materiale

I dette kapitel vil jeg redegøre for de overvejelser der ligger bag udformningen af det testmateriale som undersøgelsen bygger på. Sidst i kapitlet beskrives de praktiske forhold omkring optagelsen af materialet.

7.1 Generelle overvejelser

For at gennemføre analysen på den planlagte måde er der behov for optagelser af minimalpar, sådan at man kan sammenligne forløbet af de forskellige akustiske parametre, fx kompression ($H1:H2$), i ord med og uden stød. Det vil samtidig være ideelt hvis formantværdierne er konstante igennem vokalen, dvs. at der ikke er diftongering, og det er en fordel hvis den laveste formant ($F1$) ligger noget over $H2$.

Som nævnt er stødets realisation præget af en betydelig variation der dels viser sig fra taler til taler, men som også kommer til udtryk fra gang til gang i hver enkelt talers realisation af stød, selv i de samme ord. Variationen beskrives som fri variation der til dels kobles med distinkthed, jævnfør citatet fra Grønnum og Basbøll i afsnit 3.2.6 s. 45 og Fischer-Jørgensen 1989a i afsnit 3.1.1 s. 28 og afsnit 3.1.4 s. 38, sådan at stød under kraftig emfase evt. kan realiseres som et glottalt lukke, og svagere stød har svagere tegn på uregelmæssige stemmelæbesvingninger.

Hvis man tager som udgangspunkt at det er et retvisende billede af de faktiske forhold, så vil man være frit stillet til at vælge alene at analysere stød i strukturer som man må antage giver de bedste betingelser for de planlagte analyser. Med det udgangspunkt vil det være oplagt at se på tostavellesord med lang vokal med høj åbningsgrad: Dels er det let at finde tostavellesord som danner minimalpar med og uden stød; åbne vokalkvaliteter har højtliggende $F1$ hvilket mindsker risikoen for en komplicerende interaktion mellem $F1$ og de to nederste deltoner ($H1$ og $H2$); i lange vokaler er det venteligt at den centrale del af stødet vil falde i vokalen, hvorimod det er venteligt at artikulationen af stød i stavelser med kort vokal fulgt af sonorant konsonant vil strække sig hen over overgangen mellem vokal og konsonant og dermed komplicere de spektrale analyser.

Der er dog også gode grunde til ikke at gå så skråsikkert til værks at man udelukkende ser på stød under det man antager, er de mest favorable betingelser.

Først og fremmest viser de gennemgåede resultater fra Fischer-Jørgensen og Grønnum og Basbølls undersøgelser flere eksempler på systematisk variation i forskellige aspekter af stødrealisationen. Graden af knirk lader ifølge Fischer-Jørgensen til at afhænge af graden af distinkthed, men også af vokalens åbningsgrad (Fischer-Jørgensen 1989a, p. 50). Den nøjagtige timing af evt. knirk lader ifølge Grønnum og Basbøll til at hænge sammen med vokalens åbningsgrad (Grønnum & Basbøll 2001b, p. 247). Sandsynligheden for at der forekommer det omvendte tonale forløb som Fischer-Jørgensen nævner, ser ud til at hænge sammen med vokalens

åbningsgrad, men ser samtidig ud til at hænge sammen med talerens regionale baggrund. Som vi så, fandt Grønnum og Basbøll at varigheden af vokaler og konsonanter med stød påvirkes forskelligt af om støddordet har fokus eller ej, og varigheden af konsonanter med stød varierer betydeligt afhængig af om den stødte konsonant står ordmedialt, ordfinalt eller ytringsfinalt. Fischer-Jørgensen nævner også mulige forskelle mellem stød i en- og tostavelsesord med hensyn til f_0 og luftstrøm (Fischer-Jørgensen 1989a, p. 128). Der er altså ikke grund til at tro at realisationen af stød skulle være helt uafhængig af den fonetiske kontekst.

I og med at der ses på nye aspekter af stødets realisation ved hjælp af akustiske metoder som ikke tidligere er afprøvet på stød, er det også muligt at der vil vise sig nye tegn på variation som måske ikke er så fri. Det er fx oplagt at forestille sig at det kunne spille en rolle for udfoldelsen af det tænkte stemmeklangersforløb hvor lange stemte stræk et konkret stød er omgivet af: Er en vokal med stød omgivet af ustemte konsonanter, må klangforløbet i sagens natur afvikles helt inden for vokalens varighed, hvorimod de formodede ændringer i stemmekvaliteten potentielt kan afvikles over et længere stræk hvis en vokal med stød er omgivet af sonore lyde. Omvendt kan man også forestille sig, at der kan opstå et behov for at signal om stød ikke må fortsætte ind over en stavelsesgrænse, fx hvis et segment med stød grænser op til en betonet vokal uden stød (eksempler: "to' aser", eller "den' uge").

Som ofte er fri variation ikke fuldstændig fri: Hvor der er variation, er der næsten altid en form for systematik. Ved at undersøge et mere forskelligartet materiale kan der opstå nye muligheder for at få øje på usete og måske mere komplekse sammenhænge. I en eksplorativ undersøgelse må det derfor være formålstjenligt at kaste nettet bredt ud.

Der er altså et dobbelt hensyn at tage i udformningen af materialet. På den ene side skal det være et materiale som egner sig bedst muligt til de akustiske analyser, dvs. som tilgodeser ønskerne om minimalpar med og uden stød og en formantstruktur som intervenserer mindst mulig med bestemmelsen af styrkeforholdet mellem H1 og H2. På den anden side skal det også gerne afsøge om der er potentielle samspil med andre fonetiske eller strukturelle forhold.

7.2 Oversigt over faktorer

I det følgende vil jeg først give en oversigt over de faktorer jeg tager med, og i det efterfølgende afsnit give en uddybende begrundelse for hver. Undersøgelsen baseres på matchede ordpar med og uden stød, og materialet består derfor så vidt det er muligt af minimalpar. Da ikke alle de undersøgte strukturer giver mulighed for minimalpar, må der i visse tilfælde bruges subminimale par. Man kunne vælge at bruge nonsensord for at få absolutte minimalpar, men det er et mindre attraktivt alternativ, blandt andet fordi nogen af disse uvægerligt ville bryde med gængs dansk fonotaks.

Ud over tilstedeværelse eller fravær af stød inkluderes følgende faktorer således:

Stød på vokal eller stød på konsonant

- lang vokal
- kort vokal + sonorant konsonant

Vokalkvalitet

- i
- a

Tryk⁴⁷

- Hovedtryk
- Bitryk

Segmentel kontekst

(Ordene i parentes er eksempler på den angivne struktur⁴⁸.)

- I stavelser med hovedtryk, med eller uden stød på lang vokal, før stemt konsonant
 - $C_{\text{stemt}} + v:(') + C_{\text{stemt}}$ (li'm)
 - $C_{\text{lukke}} + v:(') + C_{\text{stemt}}$ (bar'net)
 - $C_{\text{frikativ}} + v:(') + C_{\text{stemt}}$ (fi'n)
- I stavelser med hovedtryk, med eller uden stød på lang vokal, før frikativ, lukke eller ordgrænse (markeret med #)
 - $C_{\text{sonorant}} + v:(') + C_{\text{frikativ}}$ (ma'ser)
 - $C_{\text{lukke}} + v:(') + C_{\text{lukke}}$ (ta'ber)
 - $C_{\text{lukke}} + v:(') + \#$ (ta'g)
- I stavelser med hovedtryk, med eller uden stød på stemt konsonant
 - $C_{\text{stemt}} + v + C_{\text{stemt}}(')$ (lamm'et)
 - $C_{\text{lukke}} + v + C_{\text{stemt}}(')$ (ball'er)
- I stavelser med bitryk findes et mere begrænset repertoire
 - $C_{\text{stemt}} + v:(') + C_{\text{stemt}}$ (overma'ler)
 - $C_{\text{lukke}} + v:(') + C_{\text{stemt}}$ (udta'ler)
 - $C_{\text{lukke}} + v + C_{\text{stemt}}(')$ (kæmpeball'er)

47 Jeg opererer med tre grader af tryk: hovedtryk, bitryk og nultryk.

48 Følgende notation benyttes i forbindelse med testordene: Stød markeres med apostrof efter det segment hvori stødet (højest sandsynlig) falder. Formålet med notationen er dels at disambiguere de homografe ord med og uden stød, dels at eksplicitere hvor jeg forventer stød, og dermed lette læsningen for læsere som ikke har stød i deres sprog, eller som har stød med en anden distribution end den jeg beskriver. Jeg nøjes med at bruge egentlig lydskrift hvor der er særligt behov for at præcisere den konkrete udtale.

Antal stavelser

- en stavlese
- to stavelser (tryk og evt. stød i første stavelse)⁴⁹

Morfemopbygning

- monomorfematisk (a'sen)
- rod+endelse (ma'ser)
- rod+rod (bar'nsbe'n)⁵⁰

Dertil kommer noget ekstramateriale som jeg vil kommentere efter gennemgangen af faktorerne

- læ'ner, æn'der, læsser, esser
- fil'm, mal'm, hal'm, hal'
- fi's, fir's, fa'gs, far's

⁴⁹ Det afgørende er formodentlig ikke om ordet har en eller to stavelser, men om den betonedede stavelse er den sidste eller den næstsidste i ordet. Grønnum og Basbøll benytter termerne oxyton og paroxyton om ord med tryk på sidste henholdsvis næstsidste stavelse. I Poul Andersens terminologi benævnes disse henholdsvis alfa- og betaord. I dette materiale har alle tostavelsesordene tryk på næstsidste stavelse og er dermed paroxytone/betaord.

⁵⁰ Bemærk, det er førsteleddet der analyseres.

Oversigt over ordmaterialet			
Tryk	Segment		
	[ɛæɑ]	[i]	[ɪm]
Hovedtryk	lar'men lar'met lar'm lar'me ma'ler maler ma'l male læ'ner bar'net barnlig bar'nsben barneben ma'ser mases læ'ser læser læsser esser fa'gs far's bar'sel varsel a'sen vasen aser ta'ber taber ta'g tage	li'men limet smi'ler smilet smi'l smile fi'n fine ti' tie fi's fir's	lamm'et lammet ball'er baller hal' hal'm mal'm fil'm æn'der
Bitryk	overma'ler overmaler overma'le overma'l udta'ler udtaler		kæmpeball'er kæmpeballe

Tabel 7.1. Oversigt over ordmaterialet ordnet efter stødbærende segment og trykgrad. Åben vokal, hovedtryk: lys grøn. Åben vokal, bitryk: mørkere grøn. Snæver vokal: gul. Sonorant: blå.

Oversigt over ordmaterialet						
Foregående segment	Efterfølgende segment					Tryk
	_vokal	_sonorant	_frikativ	_klusil	_# (final)	
vokal_	lamm'et lamm ^{et} ball'er ball ^{er} æn'der	fil'm hal'm mal'm			hal'	hoved-tryk
	kæmpeball'er kæmpeball ^{er}					bitryk
sonorant_		lar'men lar'm ^{en} lar'm lar'm ^e ma'ler ma'ler ma'l ma'l ^e læ'ner li'men li'm ^{en} smi'ler smi'ler smi'l smi'l ^e smile	ma'ser ma's ^{er} læ'ser læ'ser læsser læsser			hoved-tryk
		overma'ler overma'ler overma'le overma'l				bitryk
frikativ_		fi'n fi ⁿ fine	fa'gs far's fi's fir's			hoved-tryk
						bitryk
klusil_		bar'net bar'n ^{ig} bar'nsben bar'neben	bar'sel bar's ^{el}	ta'ber ta'ber	ta'g ta'g ^e ti' ti' ^e tie	hoved-tryk
		udta'ler udta'ler				bitryk
#_ (initial)			a'sen a's ^{en} vasen vas ^{en} aser			hoved-tryk
						bitryk

Tabel 7.2. Oversigt over ordmaterialet ordnet efter omgivende segmenter. Farver; se tabel 7.1.

7.3 Gennemgang af faktorer

7.3.1 Vokal vs. konsonant

Bortset fra varighedsforholdene (se afsnit 3.1.1 og 3.2) er der ikke hidtil fundet systematisk forskel på realisationen af stød i vokal og konsonant, men formantstrukturen vil være markant forskellig og vil derfor interagere med overtonemålene på forskellig vis. Derfor er det ikke utænkeligt at der kan være forskel på hvor tydeligt de forandringer i stemmekvalitet jeg leder efter, træder frem i de to tilfælde.

7.3.2 Vokalkvalitet

Som tidligere diskuteret (5.4) komplicerer talekanalens filterfunktion målingerne af stemmekildens karakteristika, og forskellige vokalkvaliteter (med hver deres karakteristiske formantværdier⁵¹) vil indvirke på forskellig vis og i forskellig grad på de foreslåede akustiske mål. Det giver som nævnt anledning til at foretrække åbne vokaler. Både Fischer-Jørgensen og Grønnum og Basbølls undersøgelser tyder dog på at vokalåbningsgrad kan spille en vis rolle for hvordan stødet realiseres. Fischer-Jørgensen omtaler en sammenhæng mellem vokalens åbningsgrad og andelen af tokens med uregemæssige svingninger (Fischer-Jørgensen 1989a, p. 50), og Grønnum og Basbøll viser en vis sammenhæng mellem vokalens åbningsgrad og timing af evt. uregemæssige svingninger (Grønnum & Basbøll 2001b, p. 247). Desuden er det påfaldende at i de dialekter som har alternativer til egentligt stød, som fx klusilspring, optræder disse kun/hovedsageligt i forbindelse med snævre vokaler. Det kan være en fonotaktisk tilfældighed, men ofte er distributionelle mønstre fonetisk motiveret (artikulatorisk, akustisk eller perceptuelt).

7.3.3 Segmentelle omgivelser

Overordnet set er målet med at variere den segmentelle kontekst at undersøge om det spiller en rolle for stødets realisering om de omgivende segmenter er stemte eller ej. Hvis min hypotese om stødet holder, er det tænkeligt at stemmekvalitetsgestussen kan have et domæne som er større end det vi normalt tilskriver stødet. Derfor kunne man tænke sig at, det gjorde en forskel hvor meget eller hvor lidt stemt materiale der omgiver stødet: Er der ustemte konsonanter både før og efter, kan stemmekvalitetsgestussen naturligvis kun udfolde sig inden for vokalen, hvorimod den kan tænkes at brede sig over et større tidsdomæne hvis muligheden er der. For de ustemte konsonanters vedkommende skelner jeg desuden mellem frikativ og lukke. Det gør jeg med tanke på at det kunne spille en rolle om der skal opretholdes en luftstrøm gennem glottis i segmentet umiddelbart efter stødet eller ej.

7.3.4 Ordgrænse og åben vs. lukket stavelse

Ordgrænse optræder i skemaet som en kontekst selvom det naturligvis ikke kan betragtes som en segmentel kontekst.

51 Som i øvrigt oftest vil variere i større eller mindre grad i løbet af vokalens varighed.

For de to enstavede støddord i hovedskemaet (ta'g, ti') gør det sig desuden gældende, at stødet falder i en fonologisk åben stavelse⁵². Herved adskiller den fonologiske stavelsesstruktur i den betonedede stavelse sig fra de øvrige testord som alle må regnes som lukkede stavelser.

Der er desuden grund til at antage at en stavelsesgrænse som samtidig er en ordgrænse, er stærkere end en tilsvarende ordintern stavelsesgrænse. Stødet i ta'g, ti', og hal' vil derfor, uanset fortolkningen af de øvrige stavelsesgrænser i materialet, efterfølges af en stærkere stavelsesgrænse.

7.3.5 en- og to-stavelsesord

Ord med stød umiddelbart før en ordgrænse må nødvendigvis have tryk på sidste stavelse, og i dette materiale er de alle enstavelsesord. I tostavelsesord med tryk på næstsidste stavelse (fx læser) optræder stødet naturligvis ikke umiddelbart før en ordgrænse. For at kunne skelne mellem en evt. betydning af stavelsesantal og stødets placering i forhold til ordgrænse medtages nogle enstavelsesord hvor stødet forekommer ord- og stavelsesinternt: må:'l, lar'm.

7.3.6 Hovedtryk vs. bitryk

Der er to grunde til at se på stød i hovedtryk i forhold til bitryksstavelser. Den ene har med distinkthed (eller prominens) at gøre, og den anden har med tonale forhold at gøre.

Distinkthed

Man regner almindeligvis med at stavelser med hovedtryk realiseres mere distinkt end stavelser med lavere grad af tryk. Observationer i Fischer-Jørgensen 1989a tyder også på at stød realiseres mindre energisk i stavelser med bitryk. Jeg venter også at se lidt svagere udsving i de akustiske mål for stemmekvalitet i forbindelse med stød i bitryksstavelser. Hvis man skulle følge Fischer-Jørgensen i at tilstedeværelse af knirk i forbindelse med stød var et spørgsmål om distinkthed, så betyder det også at man måtte forvente at se en større sandsynlighed for at stød i stavelser med bitryk realiseredes uden knirk.

Trykgruppemønsteret – det tonale fundament

Den anden grund til at lave sammenligningen er de tonale forskelle der er mellem hoved og bitryksstavelser. Som bekendt initierer hovedtryksstavelser et særligt tonalt forløb kaldet trykgruppemønsteret. Den mest dynamiske del af trykgruppemønsteret udfolder sig i hovedtryksstavelsen og den første efterfølgende ubetonede stavelse, hvorimod det tonale forløb er mere jævnt (men faldende) i efterfølgende ubetonede stavelser, jævnfør Grønnum 1992. Stød i bitryksstavelser udfolder sig altså i en tonalt

52 Den fonologiske stavelsesgrænse trækkes altid så langt til venstre som fonotaksen tillader i stavelser med fuldvokal, men så langt til højre som fonotaksen tillader før schwa. Med andre ord: Flest mulig segmenter hører til den betonedede vokal. Eksempelvis "tab-er", "mal-er" men "to-bak" og "ba-llon".

set mere homogen kontekst som kan tænkes at lette sammenligningen af de akustiske forhold som har med stødet at gøre⁵³.

Tonale forskelle mellem stavelser med og uden stød

Som vi så fandt Fischer-Jørgensen at grundtonen systematisk ligger lidt højere i den første del af stavelser med stød, og at der i det hun betegner som "the stød proper", ofte forekommer enten et fald i grundtonen, uregelmæssige svingninger eller en kombination af de to. Hun fandt desuden at der med nogen variation kan forekomme et andet trykgruppemønster i stavelser med stød hvor niveauerne er vendt om sammenlignet med det typiske (moderne københavnske) trykgruppemønster, sådan at grundtonen i den betonedede stavelse ligger højere end i den efterfølgende ubetonede stavelse.

Det fald i grundtonen Fischer-Jørgensen beskriver, må tages med forbehold for den anvendte metodes nøjagtighed i forbindelse med signaler med uregelmæssigheder i periodiciteten (hvad enten der er tale om diplofoni eller egentlig knirk) og afhænger også af det teoretiske standpunkt til begrebet grundtone i forbindelse med henholdsvis diplofoni og knirk.

Hvad sker der i bitryksstavelser med stød

Sammenligningen af stød i hovedtryksstavelser og bitryksstavelser binder desuden an til spørgsmålet om hvorvidt, eller måske snarere i hvilket omfang, stød er, eller kan være, karakteriseret ved særlige tonale forhold. Tidligere undersøgelser der har berørt emnet tonale forhold i forbindelse med stød, har så vidt jeg ved, alene beskæftiget sig med stød i stavelser med hovedtryk. Skemaet nedenfor opregner de fire relevante kombinationer af tryk og stød (A-D). Tidligere undersøgelser af tonale forhold gælder stort set udelukkende A vs. B.

	Stød	
	+	-
Hovedtryk	A	B
Bitryk	C	D

Det vil være interessant at se om de nævnte tonale forskelle der er iagttaget i stavelser med hovedtryk, også vil vise sig i stavelser med bitryk.

Umiddelbart vil man nok vente at se samme forskel mellem C og D som mellem A og B, eller måske en lidt mindre men kvalitativt sammenlignelig forskel i bitryksstavelserne (dvs. C vs. D).

Hvis de tonale forhold der vedrører stød, også er superpositioneret på andre tonale forløb (evt. med et større tidsdomæne) sådan som det er beskrevet for andre tonale komponenter i dansk (se Grønnum 1992), så er det tænkeligt at der vil vise sig en

53 Det er klart at hvis der viser sig at være en betydelig forskel i styrke eller distinkthed mellem stød i hovedtryksstavelser og bitryksstavelser, så mindskes værdien af denne sammenligning.

tydeligere tonal forskel mellem stavelser med eller uden stød når disse har bitryk. Det kunne forklares ved at det dynamiske tonale forløb der karakteriserer trykgruppens to første stavelser, "overdøver" et mere afdæmpet tonalt forløb som måtte have med stødet at gøre. Hvis det er tilfældet, kan det også tænkes at det tonale mønster vil præsentere sig som distinkt anderledes end det fremstår i stavelser med hovedtryk, netop fordi bitryksstavelser normalt er karakteriseret ved et enklere tonalt forløb. I hvilket omfang tonale forløb som har med stød at gøre og som måtte forekomme i både hoved og bitryksstavelser, skulle regnes som ens eller forskellige, må man forsøge at afgøre under skyldig hensyntagen til den tonale forskel der er mellem hoved- og bitryksstavelser i øvrigt.

Endelig kan man forestille sig at der kun ses en tonal forskel mellem stavelser med eller uden stød når disse har hovedtryk, dvs. $A \leftrightarrow B$, men $C = D$. Det ville betyde at man ikke kan betragte de tonale forskelle der er iagttaget som automatiske følgefænomener, men at man måske snarere måtte se dem som en slags ledsagende tonal accent.

7.3.7 Morfologisk struktur

Stødets distribution kan ikke beskrives uden at inddrage morfologien, se fx Basbøll 1985, 1988 og 1998). Der er ikke kendt viden om at morfologien skulle spille en rolle for hvordan stød realiseres, men det er så vidt jeg ved, ikke undersøgt, og den utvetydige kobling mellem stød og morfologi giver anledning til en vis nysgerrighed. Ved at medtage materiale med forskellig morfologisk opbygning opstår i det mindste muligheden for at se på om det kunne spille en rolle.

Konkret kunne man være nysgerrig efter at finde ud af om morfemgrænser kunne have en indflydelse på den tidsmæssige udstrækning af den foreslåede stemmekvalitetsgestus. Materialet indeholder derfor nogle tostavelsesord med tryk og stød på første stavelse og hvor den efterfølgende stavelse enten hører til samme morfem, er et selvstændigt morfem i form af en endelse eller i form af en anden rod. I de to første tilfælde er den anden stavelse ubetonet, i det sidste tilfælde har den bitryk (og i det konkrete ord desuden også stød).

monomorfematisk	(a'sen)
rod+endelse	(ma'ser)(bar'net)
rod+rod	(bar'nsbe'n)

Overvejelserne bag de valgte ord er i tråd med argumenterne for at se på stød i forskellig segmentel kontekst: Hvis stød er et dynamisk stemmekvalitetsforløb, kan krav som de omgivende elementer stiller, så spille en rolle for hvordan stød kan udfolde sig i tid? Man kunne her tænke sig at det var mere eller mindre vigtigt at holde den anden stavelse klart adskilt fra signaler om stød afhængig af den anden stavelses morfologiske status.

7.3.8 Interaktioner

Der er nogle interaktioner i materialet som man naturligvis er nødt til at have sig for øje i forbindelse med analyserne. I materialet er stavelsesstruktur, morfemkompleksitet, stavelsesantal og trykfordeling vævet ind i hinanden:

- Alle ord i materialet med tryk på sidste stavelse er enstavelsesord.
- Ord med stød umiddelbart før en ordgrænse vil naturligvis også have tryk på sidste stavelse. Det følger heraf at alle ord i materialet med stød umiddelbart før en ordgrænse også er enstavelsesord.
- Enstavelsesordene i materialet er morfologisk simple. Der er medtaget to monomorfematiske tostavelsesord med stød.
- Stød i bitryksstavelser undersøges kun i sammensatte ord hvor andetledet har to stavelser med tryk på den næstsidste stavelse.

Faktorerne kan søges adskilt i evt. videre studier hvis det viser sig at være relevant.

7.3.9 Ekstramateriale

film, halm, malm, mellem, hal med : Ord hvor stødet falder i en konsonantgruppe, er medtaget fordi udtalen af ord af denne type ved flere lejligheder har fanget min opmærksomhed. Hvis man kan afgøre stødets placering/udstrækning i ord hvor stødet falder i en konsonantgruppe, er jeg nysgerrig efter om det betyder noget om konsonantgruppens andet element tilhører samme stavelse, næste stavelse eller eventuelt et andet ord.

fi's, fir's, fa'gs, far's : Disse ord er medtaget fordi jeg er nysgerrig efter om man kan afgøre om timingen af stødet vil være den samme i stavelser med henholdsvis monofton og diftong, og i givet fald om forskellen er knyttet til realisationen af diftongen eller om den måske er mere fonologisk. I fir's realiseres vokal + r som en diftong, hvorimod fars realiseres med en monoftong som man på grund af a'ets kvalitet kan tolke fonologisk som /a+/r/. Hvis der viser sig at være en forskel mellem fi's og fir's kan man i forlængelse heraf interessere sig for om der vil være en tilsvarende forskel mellem fa'gs og far's.

læ'ner, æn'der, læ'ser, læsser, esser : Disse ord vil evt. kunne danne basis for en sammenligning med Grønnum & Basbøll 2007.

7.4 Design – argumenter for oplæsning

Eftersom de akustiske mål jeg benytter mig af i forsøget på at kvantificere stemmekvalitet, er følsomme for samspillet mellem talekanalens varierende resonanser og deltonefrekvenserne, er det sikrest at bygge analysen på sammenligning af så vidt muligt minimale par af ord med og uden stød. Denne begrænsning betyder at jeg har vurderet at det var bedst at basere afprøvningen af metodikken på oplæst tale. Dette valg træffer jeg vel vidende at oplæsning adskiller sig hørbart fra anden tale. Vi ved ikke nok om hvori forskellen mellem oplæst tale og almindelig fri tale består, men i flere henseender er oplæst tale simplere. Til denne undersøgelse hvor sigtet i høj grad er udvikling og afprøvning af nogle akustiske analysemetoder, har jeg vurderet at det var formålstjenligt og forsvarligt at benytte oplæsning i forbindelse med tilvejebringelse af materiale.

7.5 Beskrivelse af sætningsmaterialet

Oplæsningsmaterialet består af korte sætninger med mellem fire og seks betonedede stavelser med ét testord i hver sætning. Sætningerne er opbygget sådan at testordene aldrig er placeret som hverken det første eller det sidste betonedede ord i sætningen. Sætningerne er udformet så de homografe ord disambigueres. Indholdsmæssigt udfolder alle sætningerne sig inden for det samme semantiske univers. I dette tilfælde en skole og nogle aktiviteter som kan tænkes at udfolde sig der. Hver sætning forekommer mellem fire og seks gange. En samlet oversigt over sætningerne findes i appendiks G.

7.6 Bemærkning om det undersøgte materiales omfang

Jeg har valgt at starte med at se på materiale fra en enkelt informant. Flere forhold taler for at afprøve metoderne og teste hypotesen på et mindre omfangsrigt materiale frem for med det samme at kaste sig ud i at lave et fuldskala forsøg: Metodikken er ny – dels forstået på den måde at nogle af de foreslåede mål ikke er brugt i denne sammenhæng før, dels forstået på den måde at implementeringen af målemetoderne er ny. Der er derfor en vis risiko for at metoderne ikke vil virke efter hensigten. Under alle omstændigheder vil et første forsøg med stor sandsynlighed give ideer til forbedringer af metodikken. Man må også være åben over for den mulighed at information fra en enkelt informant principielt kan forkaste hypotesen hvorved behovet for et forsøg i fuld skala falder bort.

7.7 Om optagelsen

Til det indledende forsøg blev én mandlig informant på 34 år optaget. Informanten er opvokset i Hørsholm, bosiddende i Storkøbenhavn og cand.mag i lingvistik.

Optagelsen blev foretaget i Lingvistisk laboratoriums studie på Københavns Universitet Amager. Der blev benyttet en DPA 4066 headband kondensatormikrofon med kuglekarakteristik koblet til en Sound Devices SD722 harddisk optager. Optageformatet var 48 kHz samplerate, 24 bit kvantisering, ukomprimeret WAV.

Sætningerne blev præsenteret lætlæseligt, én for én, på en computermonitor placeret i god afstand til informanten. Sætningerne blev præsenteret i en på forhånd fastlagt tilfældig rækkefølge.

8 Akustisk analyse af materialet

Præsentationen af analysen disponeres på den måde at jeg starter med at gennemgå kurveforløbene i tokens af nogle af de ordpar man må formode frembyder de mest gunstige betingelser, for at se på hvordan stemmens kompression ændrer sig i løbet af ord med stød. Derefter fortsættes med ordpar med gradvist flere komplicerende faktorer. Hensigten er at give læseren et indblik i hvordan kurveforløbene kan se ud, og hvordan de læses. Gennemgangen af eksemplerne viser at der er grund til at forholde sig til de akustisk formulerede delhypoteser hver for sig og ikke kun se på den overordnede hypotese. Der præsenteres derfor endnu en stribe eksempler som viser spændvidden i forhold til de akustiske delspørgsmål. Dernæst følger en optælling af hvordan de enkelte tokens vurderes i forhold til de akustiske delhypoteser og i forhold til den overordnede hypotese. Resultaterne er overvejende negative i forhold til den overordnede hypotese og tyder desuden på uoverensstemmelser med den bagvedliggende model (dvs. Ladefogeds hierarki).

8.1 Gennemgang af eksempler på ordene ma'ler og maler

Jævnfør argumenterne i det indledende afsnit i kapitel 7 for hvilke ord der tænkes at være idelle i forhold til analysen, vil vi starte med at se på minimalpar af tostavellesord med og uden stød med lang vokal med høj åbningsgrad (dvs. her [æ:]). Som det første eksempel vælges et ordpar der udelukkende består af stemte lyde: ma'ler vs. maler. Når vokalen med stød er omgivet af stemte segmenter, er der videst mulige rammer for at følge ændringer i kompressionen, og desuden påvirkes stemmelæbernes svingninger ikke af at svingningerne skal sættes i gang eller standses umiddelbart op til vokalen med stød.

Figur 8.2 viser kurvebladet for et typisk token af ordet ma'ler, og figur 8.1 viser kurvebladet for et typisk token af det modsvarende ord uden stød: maler. Kurveblade for alle tokens af de to ord findes i appendix H.I og H.II.

Før vi begynder at se på forløbet af de kurver der har med stemmens kompression at gøre, er det nødvendigt at se på om der er forskelle imellem de to tokens der kunne have betydning for tolkningen af kurveforløbene. F1 (se nederst i ramme III) ser ud til at ligge ved samme niveau omkring 450-500 Hz i vokalen i begge ordene. F1 følger omtrent samme forløb i begge ordene med en svag stingning i den allerførste del af vokalen og dernæst et langsomt og svagt fald henimod vokalens afslutning. Det ser ud til at førsteformantens båndbredde er lidt mindre i stødordet, tydeligst i slutningen af vokalen.

Vokalen i ma'ler ^{token 2} (figur 8.2) er noget kortere end vokalen i maler ^{token 1} (figur 8.1) – omtrent 3/4 så lang.

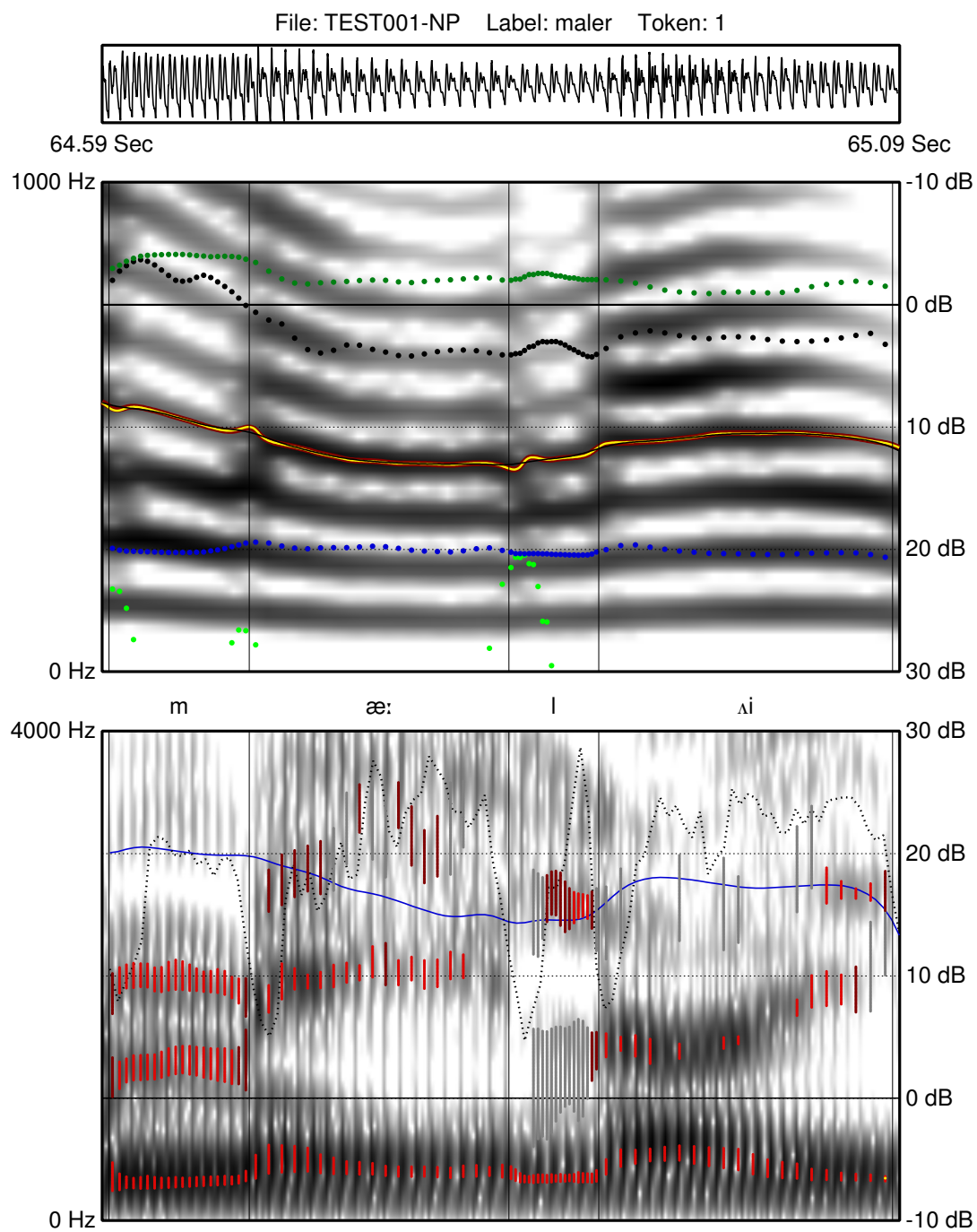
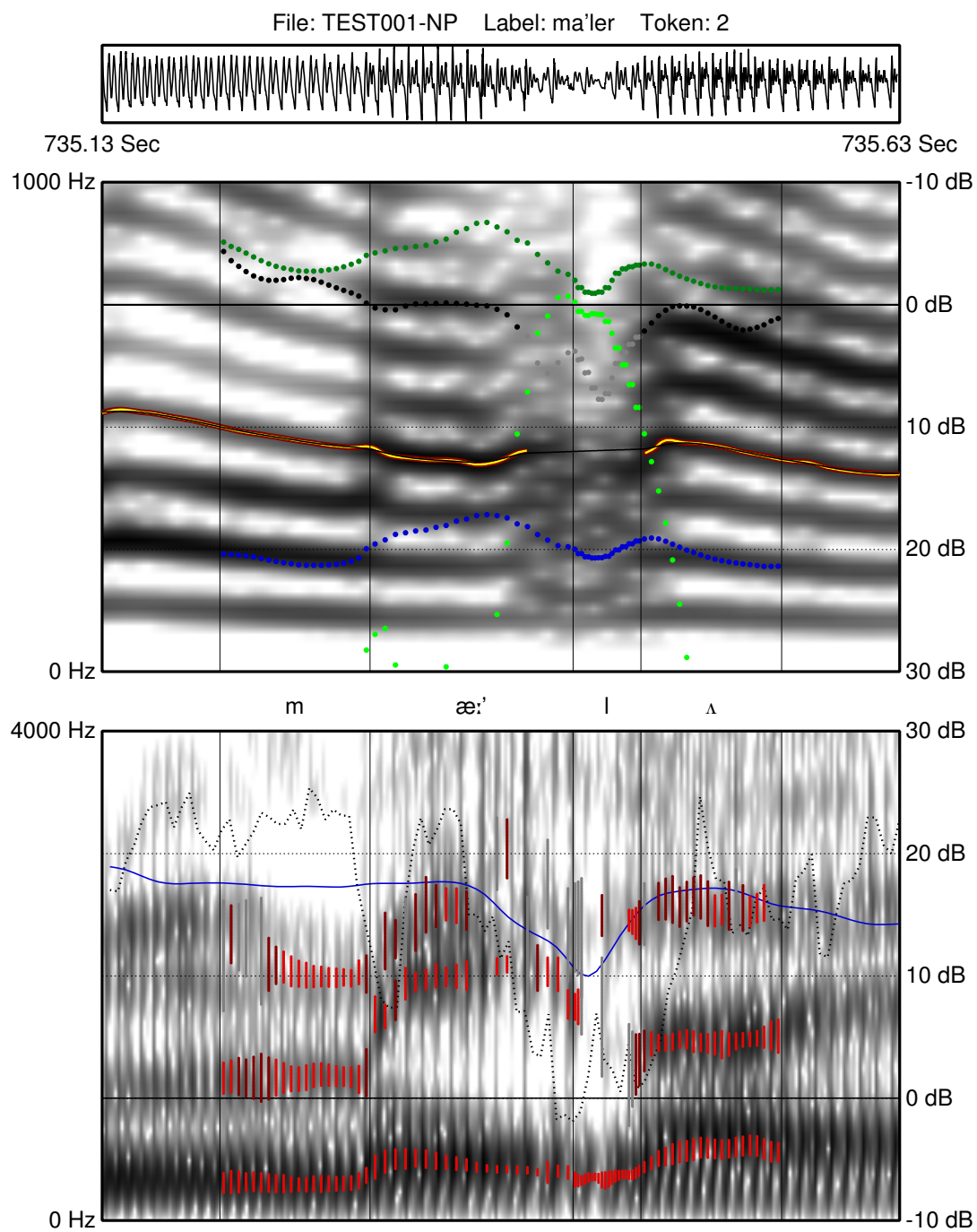


Figure 8.1. maler token nr. 1.



Figur 8.2. ma'ler token nr. 2.

Grundtonen følger i ordet uden stød det ventede faldende-stigende forløb med et minimum i sidste halvdel af [æ:] og stigning til en top i den efterfølgende ubetonede stavelse, se figur 8.1⁵⁴.

I ordet med stød udviser grundtonen et tilsvarende faldende-stigende forløb, men det ser ud til at udspille sig over et noget kortere tidsrum, og den tonale top forekommer tidligt i den efterfølgende ubetonede vokal. Samlet set ser grundtonens niveau og spændet af de tonale bevægelser dog ikke ud til at være væsensforskellige i de to ord.

I stødordet ses en lille stigning i f_0 tre fjerdedele inde i vokalen umiddelbart før svingningsmønsteret bliver diplofonisk (markeret med en sort pil i ramme II i figur 8.2). Det er ikke usædvanligt at f_0 -kurverne viser en brat op- eller nedadgående bevægelse umiddelbart før eller efter passager med diplofoni eller knirk, se fx figur 8.3 og 8.4. I en del tilfælde må det betragtes som artefakter der skyldes f_0 -algoritmens forsøg på at følge med uregelmæssighederne i svingningsmønsteret, men der er også tilfælde – som dette – hvor smalbåndssepkrogrammet viser et forløb af deltoner som ser ud til at understøtte at der reelt er tale om en kortvarigt opadgående tonal bevægelse.

De to tokens er altså ikke fuldstændigt ens hvad angår vokalens varighed, grundtoneforløb og F1. Der er dog ikke grund til at tro at de nævnte forskelle i sig selv skulle medføre væsentligt forskellige forløb for H1:H2-, BED- og CoG-kurverne.

På intensitetskurven (ramme III, blå linje) kan man se det fald eller dyk som beskrives som et at de mest konsistente akustiske træk ved stød. Jævnfør Fisher-Jørgensen (1989a) falder dykket i intensitet tidsmæssigt sammen med den glottale konstriktions maksimum.

I *maler*^{token 1} (figur 8.1) starter H1:H2 (sorte prikker) relativt højt i [m] og falder i løbet af [m]'ets sidste del og starten af [æ:]'et til et niveau hvor H1 er omkring 3 dB kraftigere end H2. Forløbet slår et lille slag i løbet af [l] og lægger sig en anelse højere i ordets sidste stavelse. BED (mørke grønne prikker) udviser et lignende forløb fra [m] til begyndelsen af [æ:], men med lidt mindre udsving. BED holder sig nogenlunde konstant gennem resten af ordets varighed. CoG (blå prikker) ligger næsten konstant på omkring 500 Hz igennem hele ordet.

Kurveforløbene ser betydeligt mere livlige ud i stødordet: H1:H2 starter i [m]'et på omtrent samme niveau som i ordet uden stød og falder også ned mod 0-linjen ved overgangen til vokalen. H1:H2-forløbet ligger forholdsvis stabilt tæt ved 0 igennem det meste af vokalens varighed, men falder forholdsvis brat umiddelbart før diplo-

54 I det konkrete token er ordets anden og ubetonede stavelse "-er" smeltet sammen med den efterfølgende ubetonede præposition "i" til en diftong der i den fonetiske analyse ikke meningsfuldt kan skilles i to stavelser. Den tonale top falder relativt sent i den dannede diftong.

fonien sætter ind. Efter den diplofoniske passage som slutter omtrent der hvor [ʌ] begynder, stiger H1:H2-forløbet igen op til nullinjen og svinger derefter lidt nedad.

BED (mørke grønne prikker) og CoG (blå prikker) udviser begge et to-toppet forløb: De starter ved samme niveau som ses i ordet uden stød, og stiger igennem vokalen indtil det højeste niveau nås omtrent 2/3 inde i vokalen. Dernæst ses et fald, eller dyk, som når sit minimum i begyndelsen af [l]. Det er sammenfaldende med det tidspunkt hvor intensitetskurven har sit minimum. Derefter stiger niveauerne igen og når deres næste top i begyndelsen af den efterfølgende vokal hvorefter begge forløb falder og flader ud igen. Man kunne få den tanke at BED- og CoG-forløbene begge kunne beskrives som en kombination af et stigende-faldende forløb der strækker sig fra den betonedede vokals begyndelse og frem til starten af den ubetonede vokal, overlejret med et faldende-stigende forløb der udspiller sig over et noget kortere tidsdomæne. I dette tilfælde strækker det faldende-stigende forløb sig fra midt i den betonedede vokal til midt i [l].

Ud over at alle tre forløb (H1:H2, BED og CoG) udviser betydeligt mere dynamiske forløb i ordet med stød sammenlignet med ordet uden stød, så ligger kurveforløbene også generelt højere fra den betonedede vokals begyndelse og indtil et stykke inde i den efterfølgende ubetonede vokal. Dog med den undtagelse at BED og CoG i forbindelse med det dyk i forløbet der ses i [l]'et, kortvarigt når ned til eller måske lidt under det niveau der ses i ordet uden stød, og selvfølgelig med det forbehold at H1:H2 ikke meningsfyldt kan bestemmes i den del hvor svingningsmønsteret er diplofonisk.

Samlet set tyder kurveforløbene på en mere komprimeret stemmekvalitet i ordet med stød allerede fra den betonedede vokals begyndelse, men tilsyneladende er kompressionen højest midt i vokalen og ikke der hvor man pr. tradition ville forvente at finde stødets "kerne", nemlig i den betonedede vokals anden halvdel (hvor intensitetsdykket er). Det er overraskende at kompressionen ser ud til at være kraftigst på et tidspunkt hvor svingningsmønsteret er jævnt periodisk, og at den falder før diplofonien sætter ind. Det er yderligere overraskende at kompressionen ser ud til at være lavest netop der hvor intensitetskurven har sit minimum, og dermed indikerer at den glottale konstriktion er kraftigst, jfr. gennemgangen af Fischer-Jørgensens resultater i kapitel 2. (Det er samtidig det tidspunkt i stød hvor Fischer-Jørgensens forsøg med inverteret filtrering tydede på at den lukkede fase i de glottale svingningers forløb var længst (dvs. lav OQ), et træk der ellers kobles til højere kompression).

Figur 8.3 er et andet eksempel på samme ord med stød. Bedømt ud fra udseendet af spektrogrammerne ser det ikke ud til at adskille sig væsentligt fra eksemplet i figur 8.2. Men H1:H2-, BED- og CoG-kurverne følger ikke det samme forløb som vi så i figur 8.2.

Sammenlignet med eksemplet i figur 8.2 ligger både BED, CoG og navnlig H1:H2 lavere i vokalens første halvdel. H1:H2 ligger på næsten samme niveau som i ordet uden stød (figur 8.1). BED og CoG stiger ikke til helt så høje niveauer som i figur 8.2

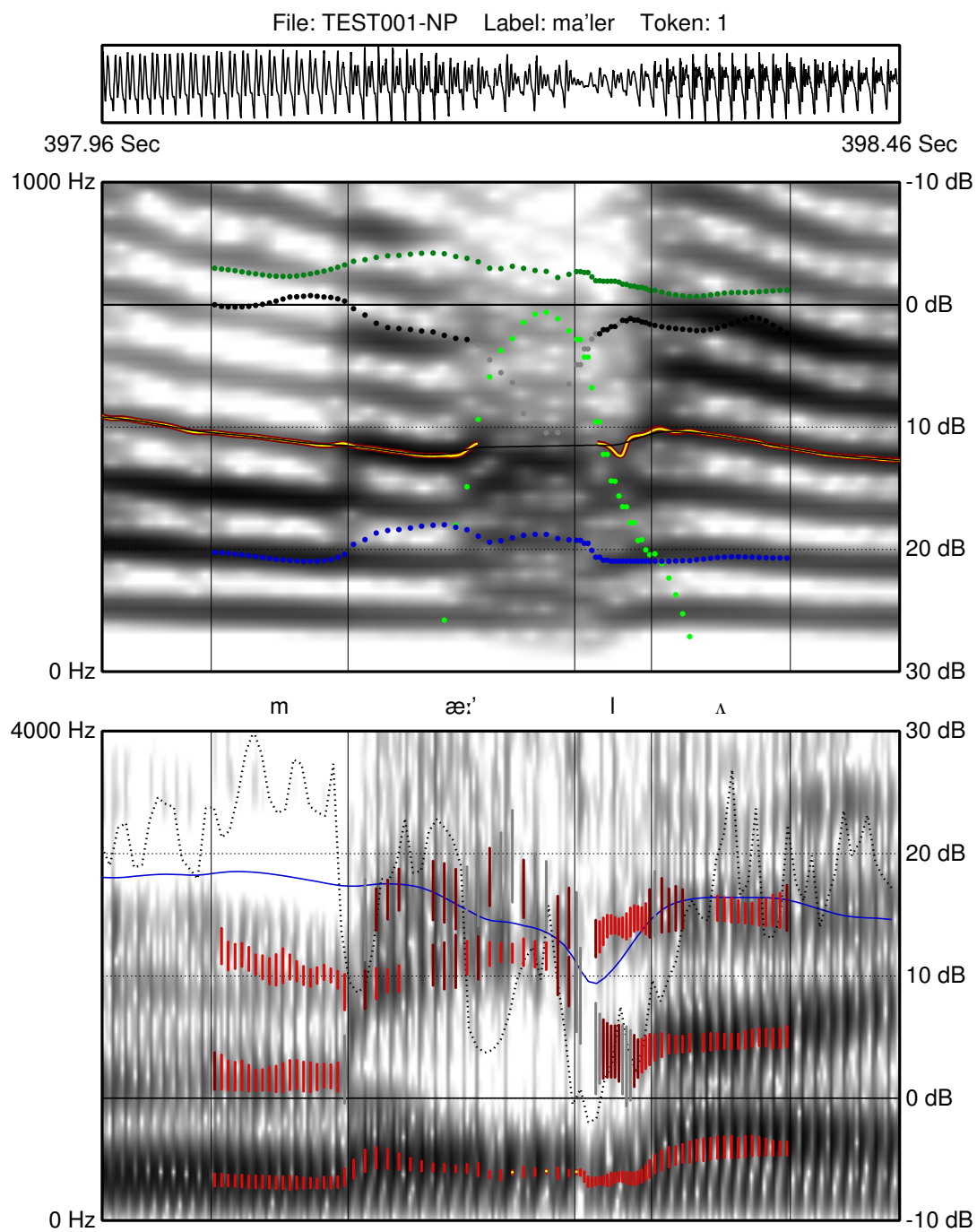
og niveauet falder allerede i [l] ned til samme niveau som ses i ordet uden stød. Hvor BED- og CoG-forløbene i figur 8.2 havde et markant dyk sammenfaldende med dykket i intensitetskurven, ser man i figur 8.3 kun antydningen af et dyk i BED- og CoG-forløbene som desuden forekommer betydeligt tidligere, nemlig tæt ved overgangen til diplofoni (tidspunktet er markeret med en blå pil). Dermed falder det ikke tidsmæssigt sammen med det dyk i intensiteten der kan ses i ramme III (markeret med en sort pil).

Den tidsmæssige afstand fra intensitetsdykket til det sted hvor kompressionen ser ud til at være på sit højeste, dvs. der hvor BED- og CoG-forløbet topper, er længere i eksemplet i figur 8.3 sammenlignet med figur 8.2.

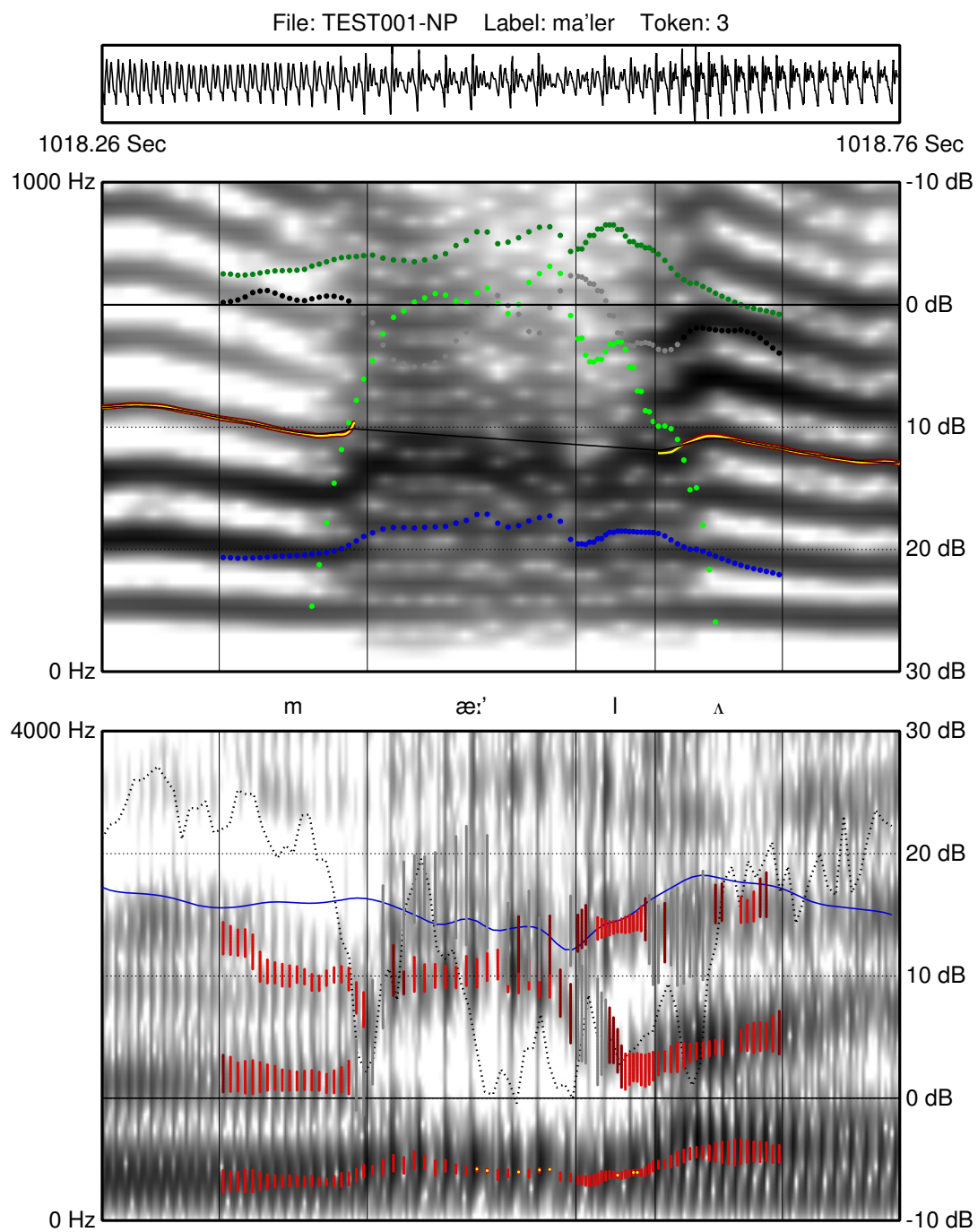
Figur 8.4 viser endnu et eksempel på samme ord med stød, men i dette tilfælde er der meget mere udbredt diplofoni og til dels knirk. (Der er passager hvor der er tydelig diplofoni, fx i vokalens første halvdel, og omkring overgangen fra [l] til [ʌ], mens den resterende del hverken er tydeligt diplofonisk eller så tydeligt kaotisk – dvs. helt uden deltonestruktur – som man normalt forventer ved knirk).

I dette tilfælde viser BED og CoG tegn på at kompressionen overordnet set udviser et stigende-faldende forløb der minder ganske meget om det vi så i figur 8.2. I dette tilfælde ses også et mindre dyk i BED og CoG der hvor intensitetskurven (blå linje i ramme III) når det laveste niveau. Eksemplet adskiller sig fra de foregående ved at BED- og CoG-forløbene er stigende og når deres højeste niveau i en passage hvor svingningsmønsteret ikke er jævnt periodisk.

I dette eksempel indikerer BED og CoG at den højeste kompression forekommer meget tæt på intensitetsdykket. Selvom det interpolerede f_0 -estimat formodentlig rammer lidt for lavt henimod slutningen af den diplofoniske passage og derfor skubber lidt til BED-værdierne, så er der generelt en overbevisende overensstemmelse mellem CoG- og BED-forløbene. Men da den interpolerede f_0 (sort linje, ramme II) nok ikke er helt korrekt, mener jeg det er mest rimeligt at bedømme hvor kompressionen er på sit højeste på baggrund af CoG-forløbet.



Figur 8.3. ma'ler token nr. 1.



Figur 8.4. ma'ler token nr. 3.

8.2 Præliminær sammenfatning på baggrund af de første eksempler

Tilsammen tegner disse fire eksempler et blandet billede af observationer i forhold til forventningerne.

Som ventet ser man i de gennemgåede eksempler på ord med stød:

- Uregelmæssige svingninger i form af diplofoni og i mindre omfang knirk i alle de tre gennemgåede eksempler på stød.
- Det karakteristiske fald eller dyk i intensiteten med et minimum tæt ved overgangen mellem den betonedede vokal og l'et ses også i alle de tre eksempler på ord med stød.
- Kompressionen ser ud til at være højere i ordene med stød end i eksempelordet uden stød. Den højere kompression ses i vokalen og [l]'et og til dels i den første del af den efterfølgende ubetonede vokal.

Men der er også overraskelser: Sammenhængen mellem disse tre forhold er ikke som ventet:

Der ser ikke ud til at være en fast kobling mellem tegn på højere kompression og forekomsten af uregelmæssige svingninger (dvs. diplofoni eller knirk). Vurderet ud fra BED og CoG kan kompressionen være omtrent lige høj og lige lav i modal som i ikke-modal stemthed. Det illustreres fx af de højtliggende BED- og CoG-værdier i figur 8.2 og 8.4, og de lavtliggende BED- og CoG-værdier i figur 8.2 og 8.1. Det er heller ikke sådan at kompressionen nødvendigvis stiger op til passager med knirk eller diplofoni.

Der ser heller ikke ud til at være en direkte kobling mellem dykket i intensitet og stemmens kompression. I et af de gennemgåede eksempler (figur 8.4) er kompressionen højest umiddelbart før (men ikke sammenfaldende med) intensitetsdykket, mens kompressionen i de andre eksempler topper tidligere i vokalen med stød. Det forhold at der kan være ret stor forskel i afstanden i tid mellem den kraftigste kompression og intensitetsdykkets minimum, og at den svageste kompression i de ikke-modale passager forekommer samtidig med intensitetsdykket, må betyde at de mekanismer der er bag den glottale konstriktion som forårsager intensitetsdykket, og de mekanismer der styrer kompressionen, ikke kan være de samme.

Intensitetsfaldet ser altså ikke ud til at være en direkte konsekvens af meget kraftig kompression sådan som det blev skitseret i forbindelse med præsentationen af hypotesen (jfr. afsnit 4.2). Der er heller ikke umiddelbart nogen andre faktorer der kan forklare hvorfor BED og CoG i 4 af de 5 tokens af "ma'ler" topper i den modale del af vokalen.

Det står altså klart at allerede resultaterne fra det første sæt af eksempler udfordrer den foreslåede hypotese. Men, én svale gør ingen sommer: Selvom de første stikprøver ser ud til at modsige hypotesen, så er det nødvendigt at se på et større

materiale før det med nogen sikkerhed kan afgøres om hypotesen eventuelt må forkastes.

8.3 ma'ser vs. mases

Det næste ordpar vi vil se på, *ma'ser* og *mases*, er ikke et perfekt minimalpar, men hvad angår den betonedede stavelse, er det rimeligt at betragte dem som minimalpar: Begge ord realiseres som tydelige tostavellesord henholdsvis [mæ:'sʌ] og [mæ:səs], og bortset fra stødet er det i den ubetonede stavelse at de adskiller sig fra hinanden.

Figur 8.5 og 8.6 viser kurveblade for to typiske eksemplarer af henholdsvis *ma'ser* og *mases*. Alle kurveblade for de to ord er gengivet i appendix H.III og H.IV.

Der ser ikke ud til at være væsentlige forskelle imellem *ma'ser* og *mases* med hensyn til f_0 -niveau og F1 i den betonedede vokal, men det er vanskeligt at sige om grundtonen som helhed forløber på samme måde i trykgruppens to første stavelser i de to ord fordi der er diplofoni eller knirk i den ubetonede stavelse i alle (!) de fem registrerede tokens af den stødløse part *mases*. I lighed med hvad vi så for *ma'ler* vs. *maler* i det foregående eksempel, er den betonedede vokal i *ma'ser* noget kortere end den betonedede vokal i de modsvarende ord uden stød; *mases*. Der er ikke grund til at tro at de forskelle der kan iagttages, vil stå i vejen for at sammenligne H1:H2-, BED- og CoG-forløbene i ordenes første stavelse.

Ordparret *ma'ser* – *mases* adskiller sig primært fra det foregående ordpar ved at vokalen efterfølges af en nominelt ustemt konsonant [s]. I distinkt udtale vil stemmelæbernes svingninger ophøre ved overgangen fra vokal til frikativ, og der vil være en jævn luftstrøm gennem glottis der bruges til at danne [s]-lydens turbulente støj. Stemmelæbernes svingning standses dog ikke brat, og stemtheden kan evt. opretholdes i [s]-lyden. I figur 8.6 og 8.7 er gengivet eksempler på tokens henholdsvis med og uden stemthed i frikativ: I *mases*^{token 3} (figur 8.6) kan det ses af både oscillogrammet og smalbåndsspektrogrammet at stemtheden opretholdes hele vejen igennem [s]-lyden, men med meget lav kompression: Der ses hovedsagelig energi ved H1 og meget lidt ved H2. Modsat i *mases*^{token 2} (figur 8.7) hvor stemtheden midt i [s] kortvarigt ophører helt.

I de to tokens af ordet *mases* der er gengivet i hhv. figur 8.6 og 8.7, indikerer H1:H2, BED og CoG samstemmende gradvist faldende kompression i forbindelse med overgangen fra den stemte vokal til den (nominelt) ustemte frikativ.

Forløbet ser ikke ud til at være kvalitativt anderledes i ordene med stød. I *ma'ser*^{token 2} (figur 8.5) indikerer BED, CoG og til dels H1:H2 (der kun kan bestemmes i de passager hvor svingningerne er regelmæssige) et tilsvarende faldende forløb op imod [s]. Det spiller tilsyneladende ikke nogen væsentlig rolle for det faldende forløb at der er stød og at der forekommer uregelmæssige svingninger netop i vokalens sidste del.

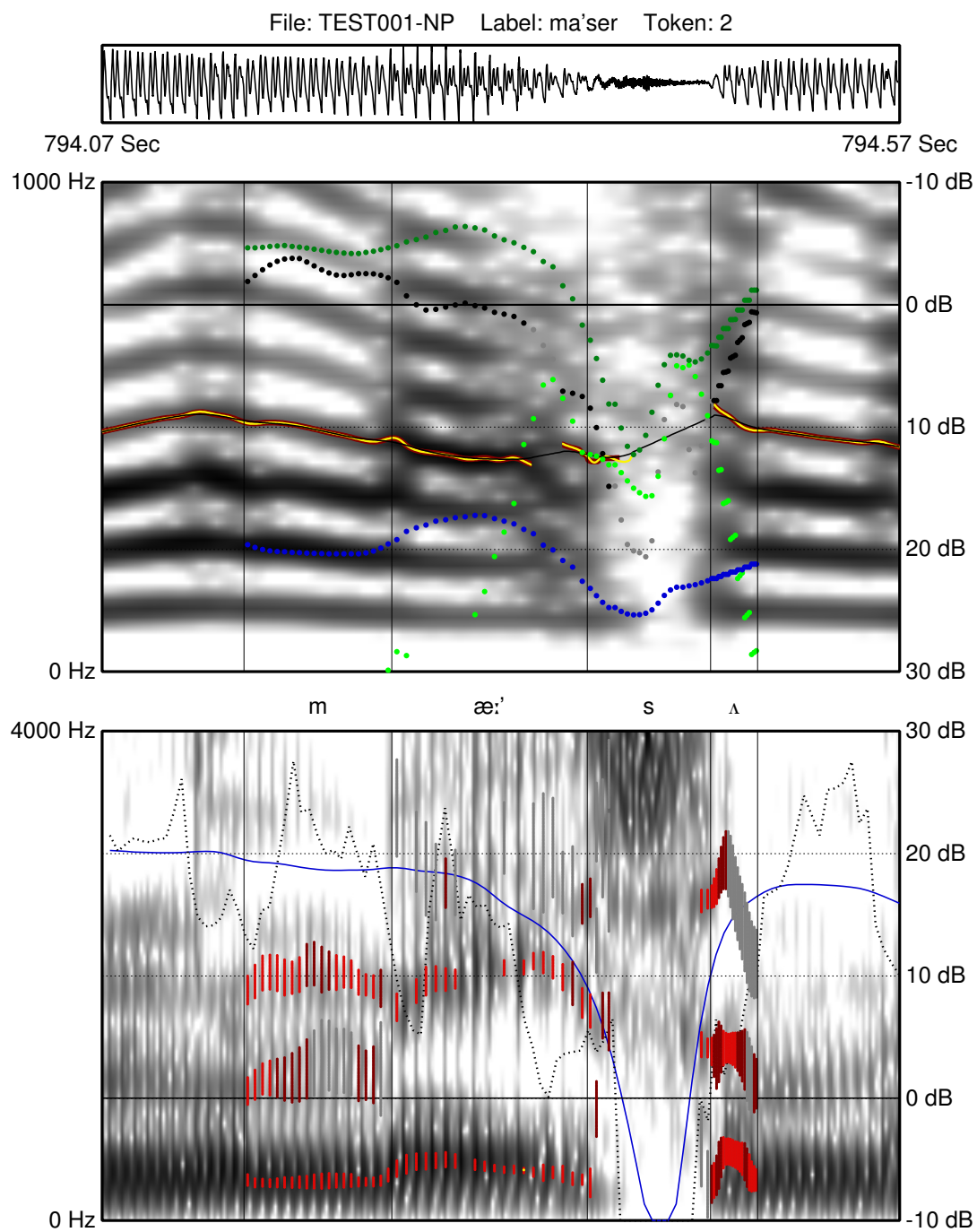


Figure 8.5. ma'ser token nr. 2.

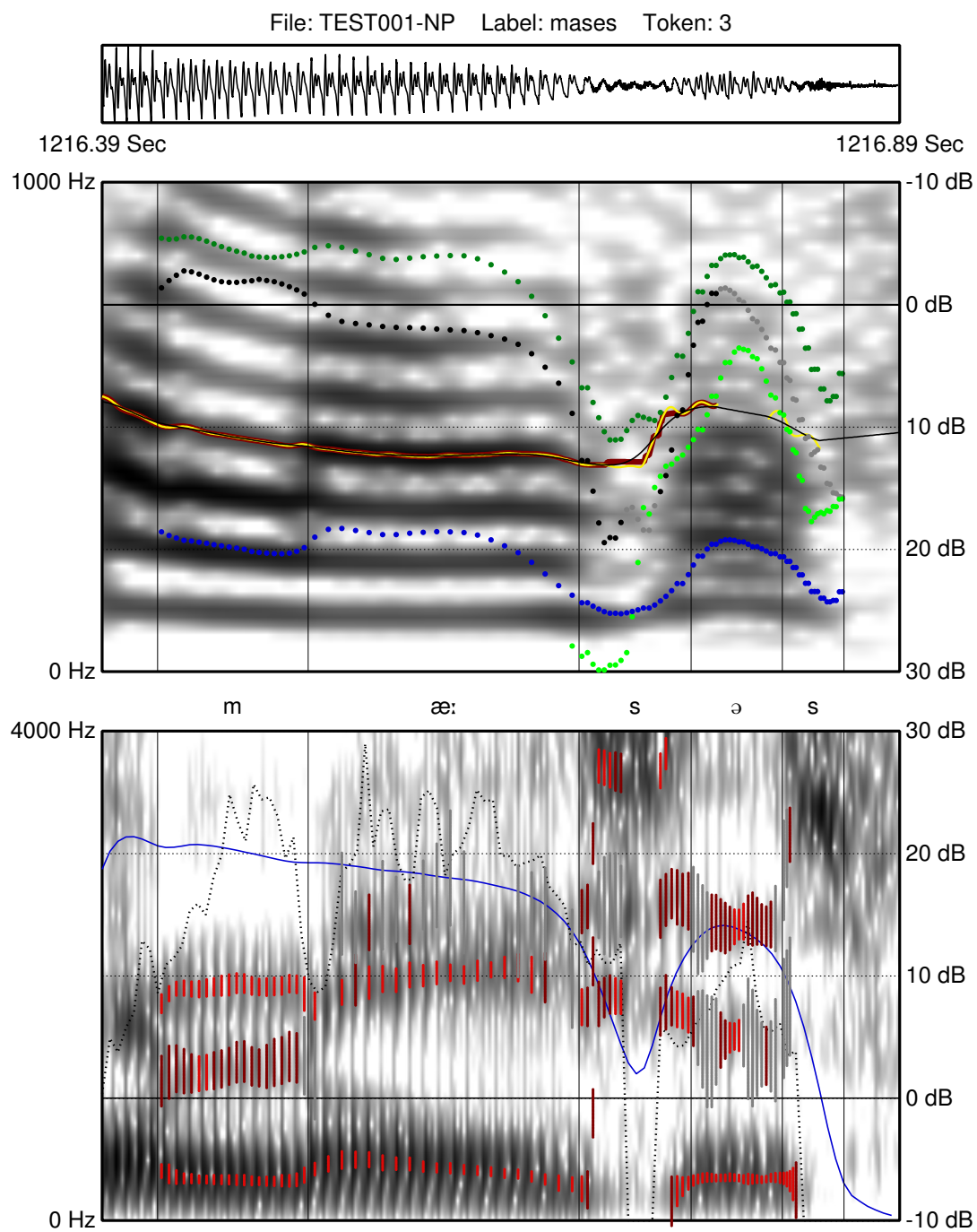


Figure 8.6. mases token nr. 3.

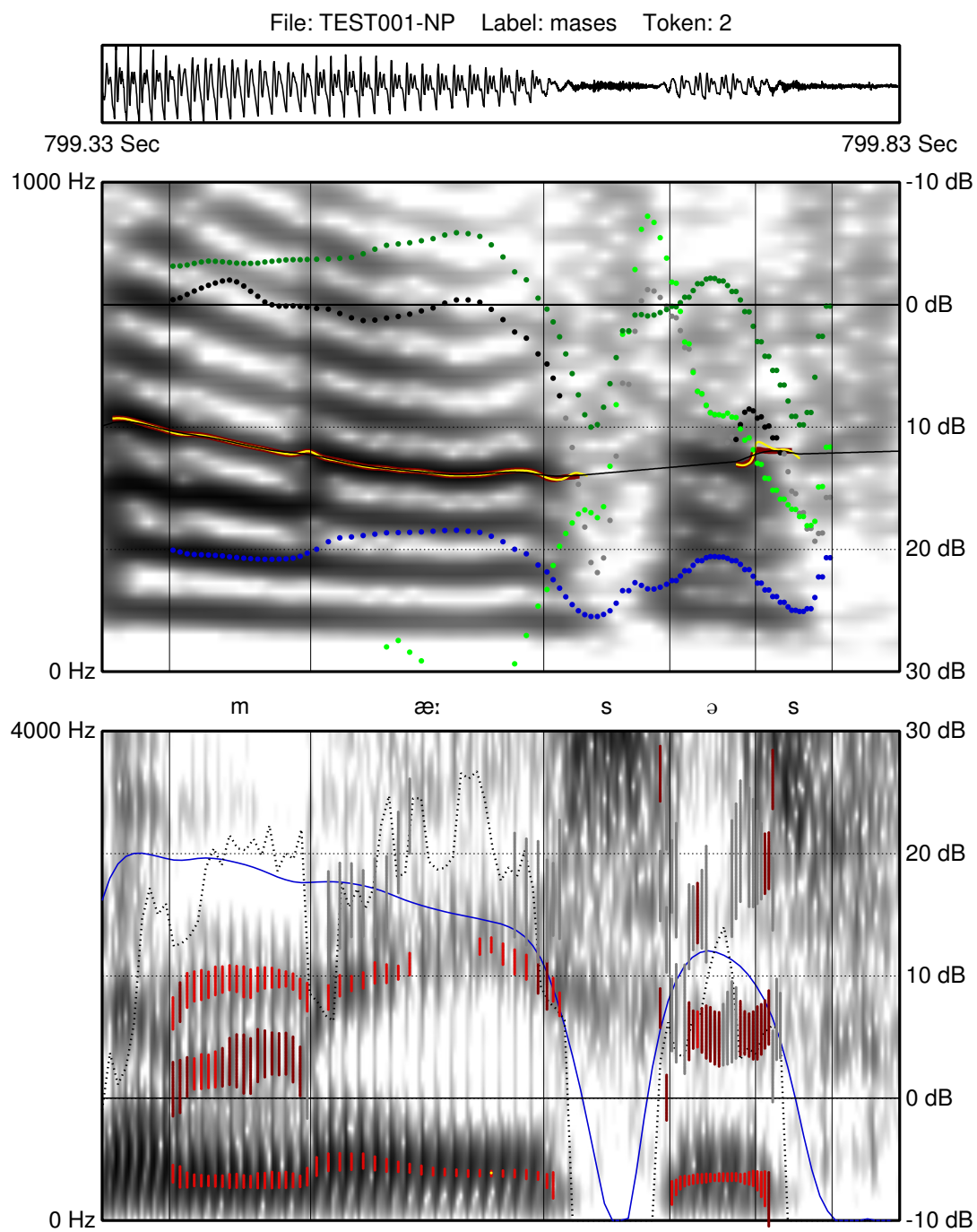


Figure 8.7. mases token nr. 2.

I alle 4 tokens af *ma'ser* ser man tegn på at kompressionen i den betonedede vokal er højest omtrent midt i vokalen (eller lidt tidligere) og før svingningsmønsteret bliver uregelmæssigt. Som man kan se på eksemplet i figur 8.5, kan det være svært at sige præcis hvor kompressionen er på sit højeste. Vurderingen afhænger af om man lægger mest vægt på H1:H2, BED eller CoG. Det er værd igen at pointere at der ikke er noget der tyder på at kompressionen skulle være højest i vokalens anden halvdel, og at der heller ikke er tegn på at kompressionen skulle være højere i fasen med uregelmæssige svingninger.

Det er dog ikke sådan at der slet ikke er forskelle imellem de to sæt af tokens: H1:H2, BED og CoG når lidt højere niveauer i *ma'ser* sammenlignet med *mases*. Forløbene er også mere dynamiske end i ordene uden stød, men forskellen er ikke så entydig som i det første ordpar vi så på (sammenlign evt. de enkelte tokens gengivet i appendix H.III og H.IV). Ser man fx på *mases*^{token 2} der er gengivet i figur 8.7, ser H1:H2 og BED ud til at udvise et forløb der når lige så højt op som de forløb der ses i *ma'ser*, og også et forløb med omtrent lige så store udsving. CoG udviser dog et mere jævnt forløb i det pågældende token så de store udsving for H1:H2 må nok tages med forbehold.

8.4 *ta'ber* vs. *taber*

Næste ordpar er *ta'ber* – *taber*. I dette eksempel er den betonedede vokal omgivet af ustemte konsonanter på begge sider, dvs. at den betonedede vokal udgør et isoleret stræk af stemthed som stødet ifølge gældende beskrivelser må udfolde sig inden for. Begge konsonanter er desuden lukkelyde hvilket betyder at luftstrømmen igennem talekanalen (nominelt set) standses helt før og efter vokalens stemte interval.

Figur 8.8 og 8.9 viser kurveblade for to typiske eksemplarer af henholdsvis *taber* og *ta'ber*. For dette ordpar ser der ud til at være lidt større forskelle hvad angår f_0 : Hvor grundtonen i ordene uden stød er svagt faldende i den betonedede vokal, ses der i ordene med stød et ret tydeligt fald i første halvdel af den betonedede vokal som muligvis flader ud eller vender til en stigning umiddelbart før svingningsmønsteret bliver uregelmæssigt. Grundtonen starter på et højere niveau i den betonedede vokal i ordene med stød og bevæger sig ned til omtrent samme niveau som i ordene uden stød. Spændet fra den laveste f_0 i den betonedede vokal til den højeste f_0 i den efterfølgende ubetonede stavelse er lidt mindre i ordene med stød end i ordene uden stød.

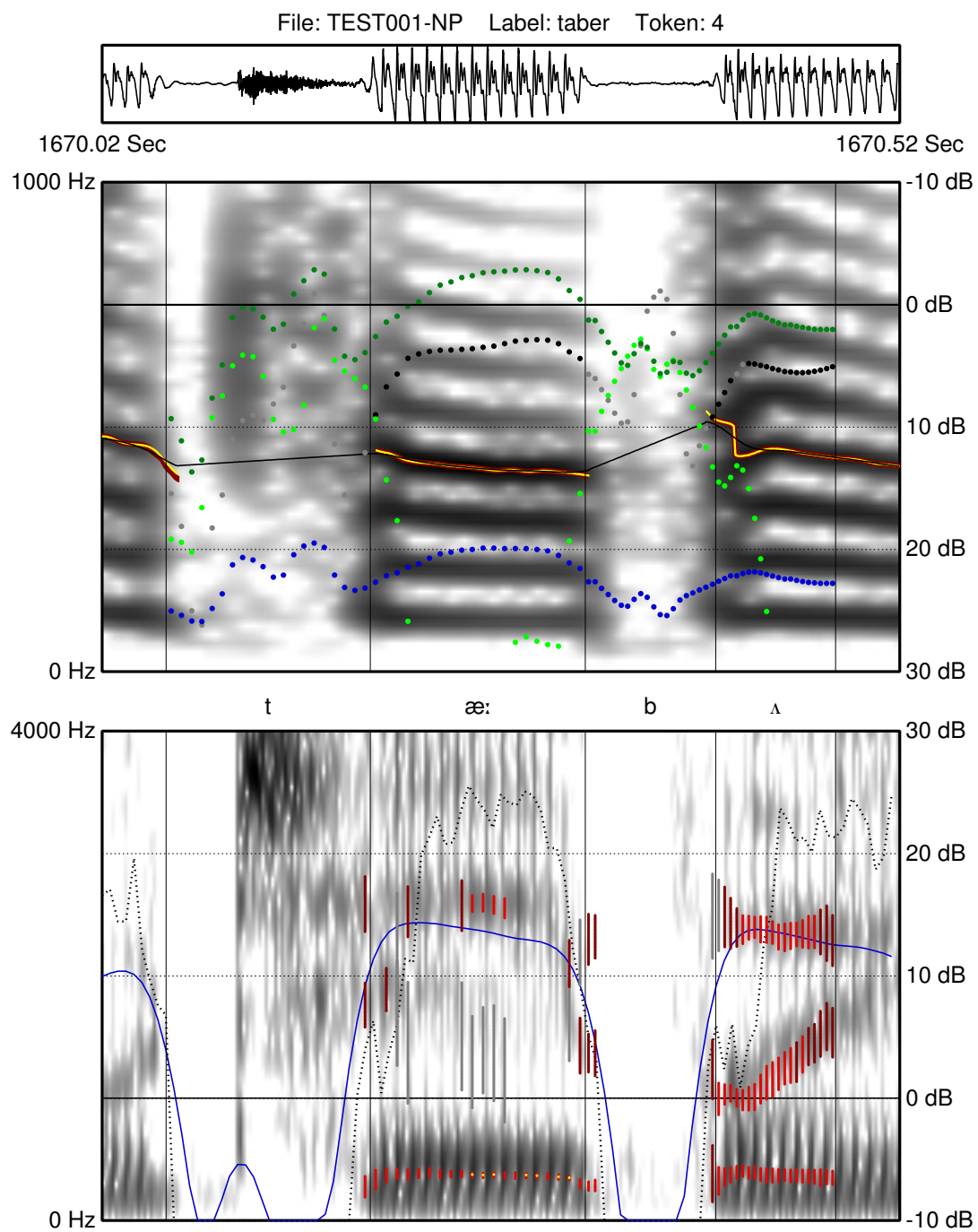
F1 ligger lidt højere i den betonedede vokal i ordene med stød. Den lidt højere F1 i ordene med stød kunne bidrage til lidt højereliggende CoG- og BED-forløb. Man kunne have forestillet sig at den højere og stejlt faldende f_0 i begyndelsen af den betonedede vokal i stødordene (se figur 8.9 og 8.10) kunne have betydet at CoG og BED (og måske H1:H2) ville have ligget højere i begyndelsen af vokalen og være faldende i takt med f_0 . Alle tre mål følger imidlertid en stejlt opadgående kurs i begyndelsen af den betonedede vokal, så hvis der er en effekt af f_0 , så overskygges den af ændringer i kompressionen.

Også i dette ordpar er den betonedede vokal i stødordene i alle tilfælde kortere end den betonedede vokal er i ordene uden stød.

I ord uden stød stiger kompressionen i vokalsens begyndelse fra et ret lavt niveau til et moderat højt niveau, og tilsvarende falder kompressionen i slutningen af vokalen før stemtheden ophører. Kompressionen er højest ca. to tredjedele inde i vokalen og lavest i begyndelsen og slutningen af vokalen hvor stemtheden henholdsvis sætter ind og standser.

I lighed med de tidligere ordpar er kompressionen lidt eller noget højere i ordene med stød. Kompressionen stiger samtidig mere stejlt i vokalsens begyndelse sammenlignet med ordene uden stød. I hovedparten af de foregående eksempler tyder analyserne på at kompressionen i vokaler med stød er højest i den modale del af vokalen. I modsætning hertil ser kompressionen her ud til at holde et højt niveau også ind i den ikke modale del: BED- og CoG-forløbet tyder på at kompressionen er næsten lige så høj i den sidste del af vokalen hvor der er uregelmæssige svingninger, og i ét eksempel (ta'ber token 5, figur 8.10) ser det ud til at kompressionen er højest i vokalsens sidste del. Kompressionen opretholdes altså på et højt niveau indtil stemtheden ophører, og ser derfor ud til at falde relativt stejlt ved vokalsens afslutning.

Forløbet i ta'ber token 5 hvor kompressionen vurderet ud fra BED og CoG ser ud til at stige helt op det tidspunkt hvor stemtheden ophører, er et (indtil nu sjældent) eksempel der kunne se ud til at være i tråd med den foreslåede model.



Figur 8.8. taber token nr. 4.

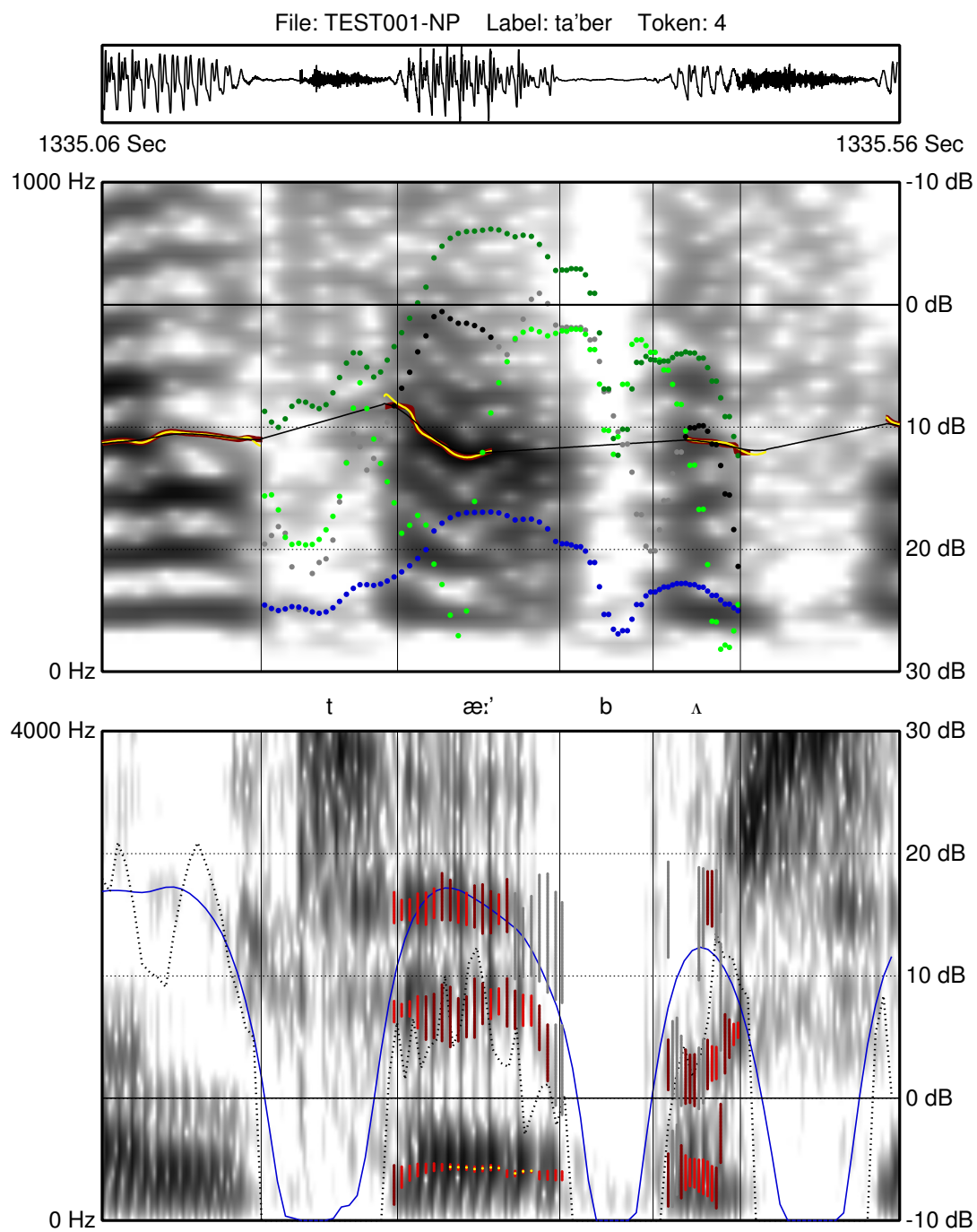
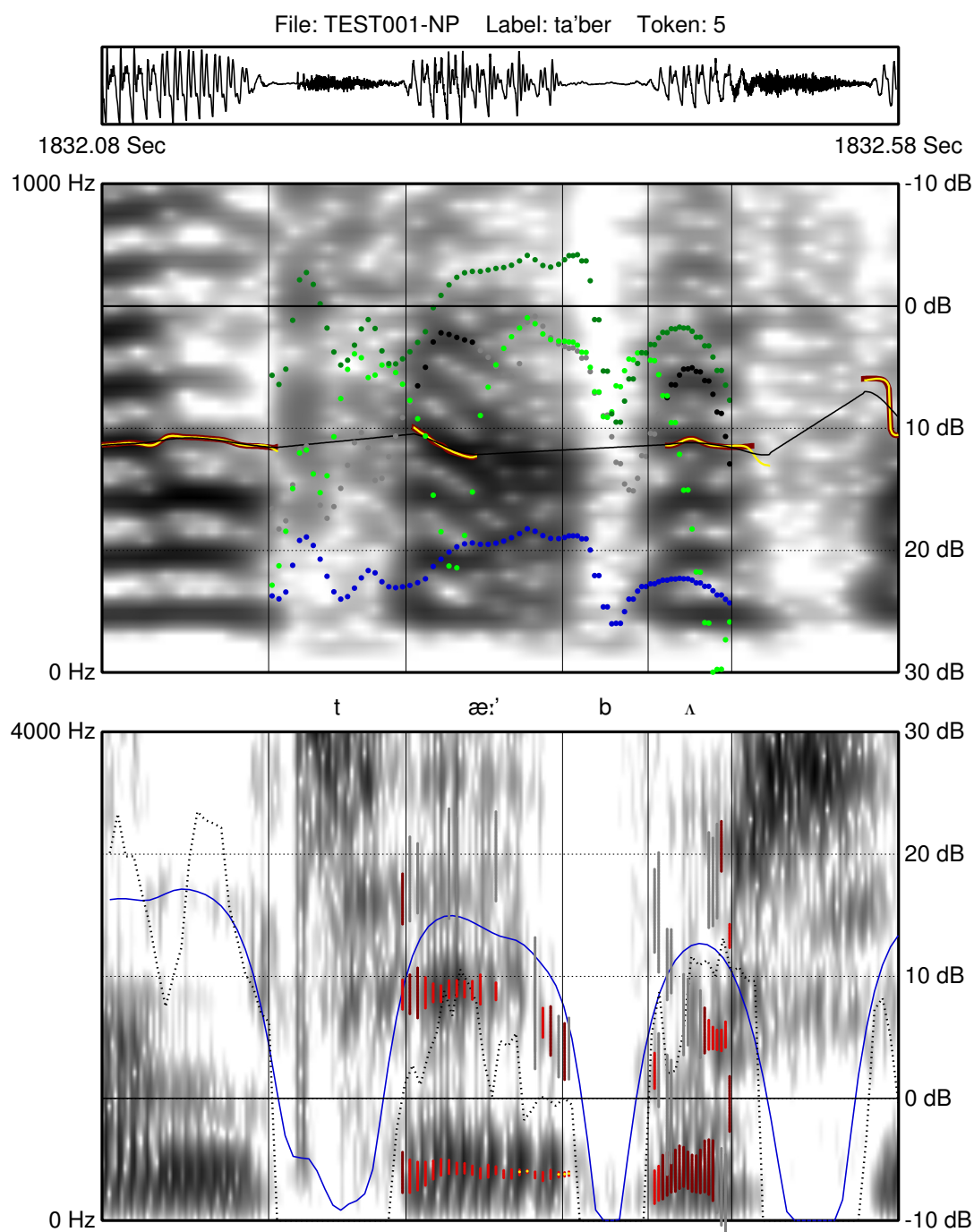


Figure 8.9. ta'ber token nr. 4.



Figur 8.10. ta'ber token nr. 5.

8.5 li'men vs. limet

I det følgende vil vi se på et ordpar med snæver vokal: limet vs. li'men. Igen er der ikke tale om et perfekt minimalpar, men bortset fra stødet er det i den ubetonede stavelse at de adskiller sig fra hinanden, og hvad angår den betonedede stavelse er det rimeligt at betragte dem som minimalpar. Figur 8.11 og 8.12 viser kurveblade for to typiske eksemplarer af henholdsvis limet og li'men.

Grundtonen følger i begge ord det ventede lav-høj forløb, men i ordene uden stød er der tale om en ret beskeden forskel fra laveste til højeste f_0 -værdi. Tager man limet^{token 2} (figur 8.11) som eksempel, har grundtonen sit minimum i slutningen af den betonedede vokal, mens den højeste grundtone ses hen imod slutningen af den ubetonede stavelse [əð] ved overgangen til det efterfølgende ord. Ser man på li'men^{token 2} som modparten med stød, ses den laveste f_0 tidligere, nemlig omtrent midtvejs til 2/3 inde i den betonedede vokal. I 3 af de 5 tokens med stød (li'men^{token 1, 2 & 5}, se appendix H.VII) er grundtonen tydeligt mere faldende i den betonedede vokal frem til det sted hvor periodiciteten bryder op, sammenlignet med ordene uden stød, se li'men^{token 2} figur 8.12 sammenlignet med limet^{token 2}, figur 8.11.

Fordi den meget lavtliggende F1 i vokalen [i] medfører at selv meget beskedne ændringer i F1 kan have stor indflydelse på H1:H2-, BED- og CoG-målene, må disse observationer dog tages med nogen forbehold.

Det kunne se ud til at H1:H2 generelt ligger lidt højere i ordene med stød, mens der ikke ser ud til at være nogen systematisk forskel for BED. CoG udviser et mere livligt forløb i stødordene; sammenlign fx li'men^{token 2 & 4} i figur 8.12 og 8.13 med limet^{token 2} i figur 8.11.

Intensitetsforløbet i ordene med stød udviser det typiske faldende forløb med et minimum tæt ved vokalens afslutning.

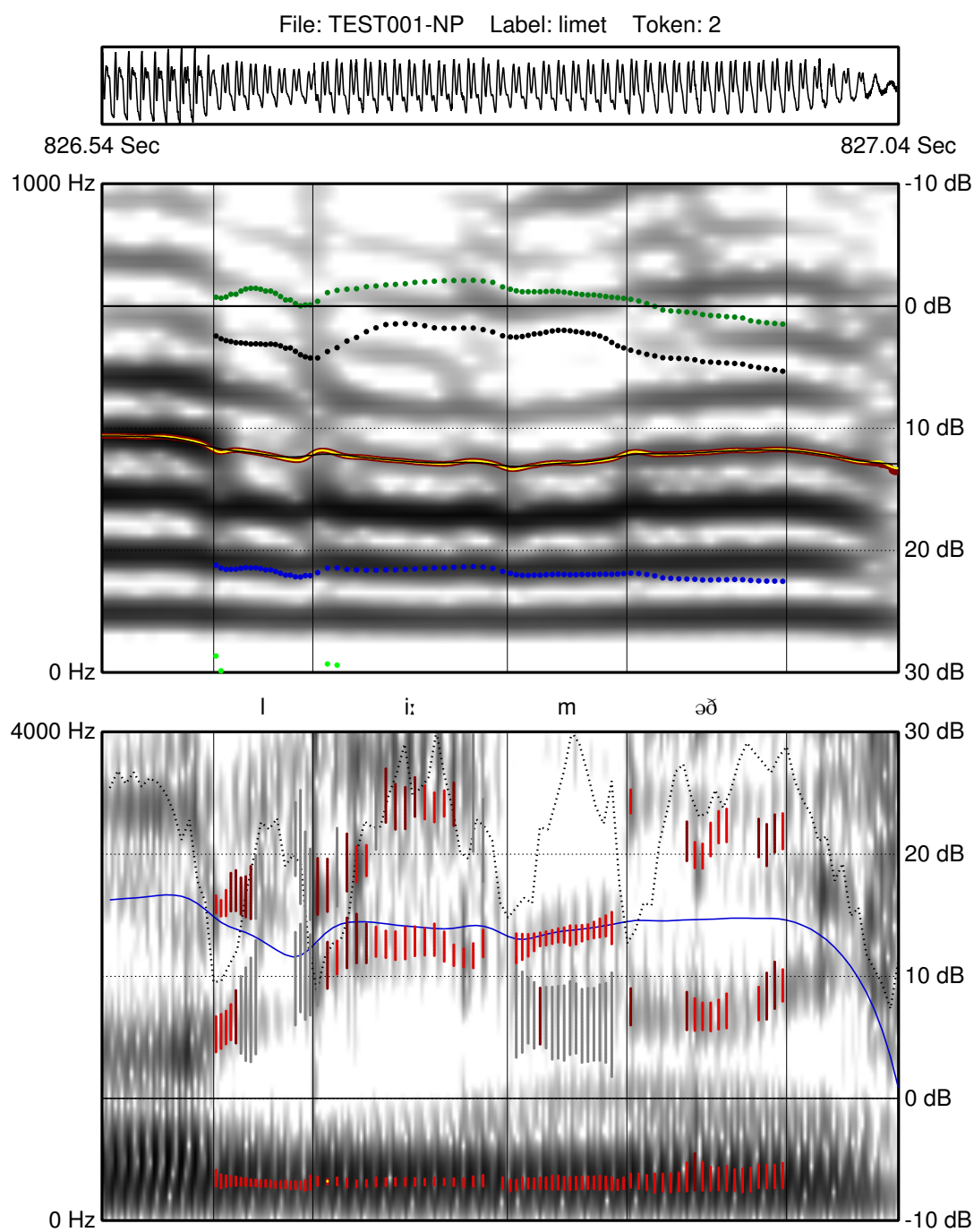
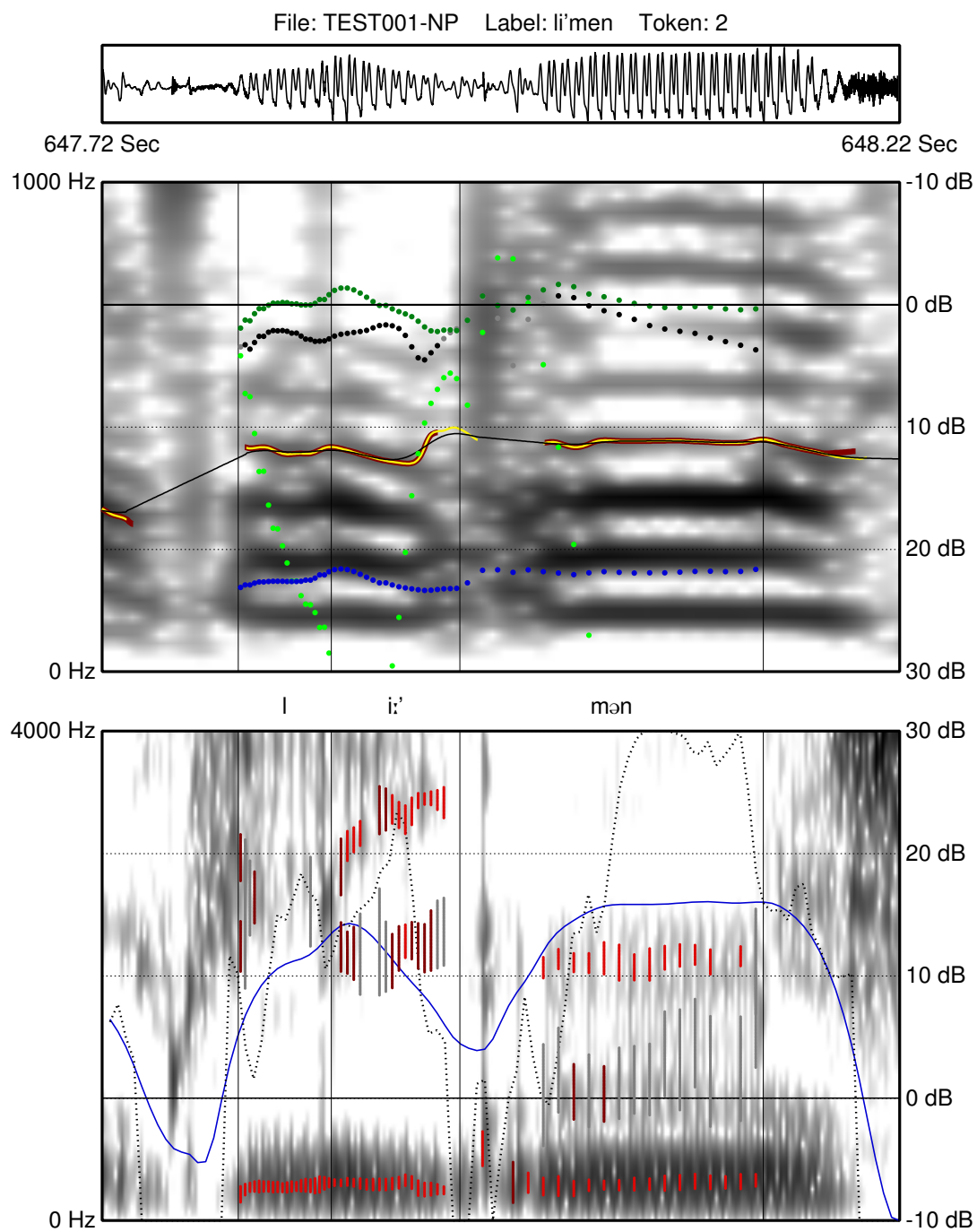
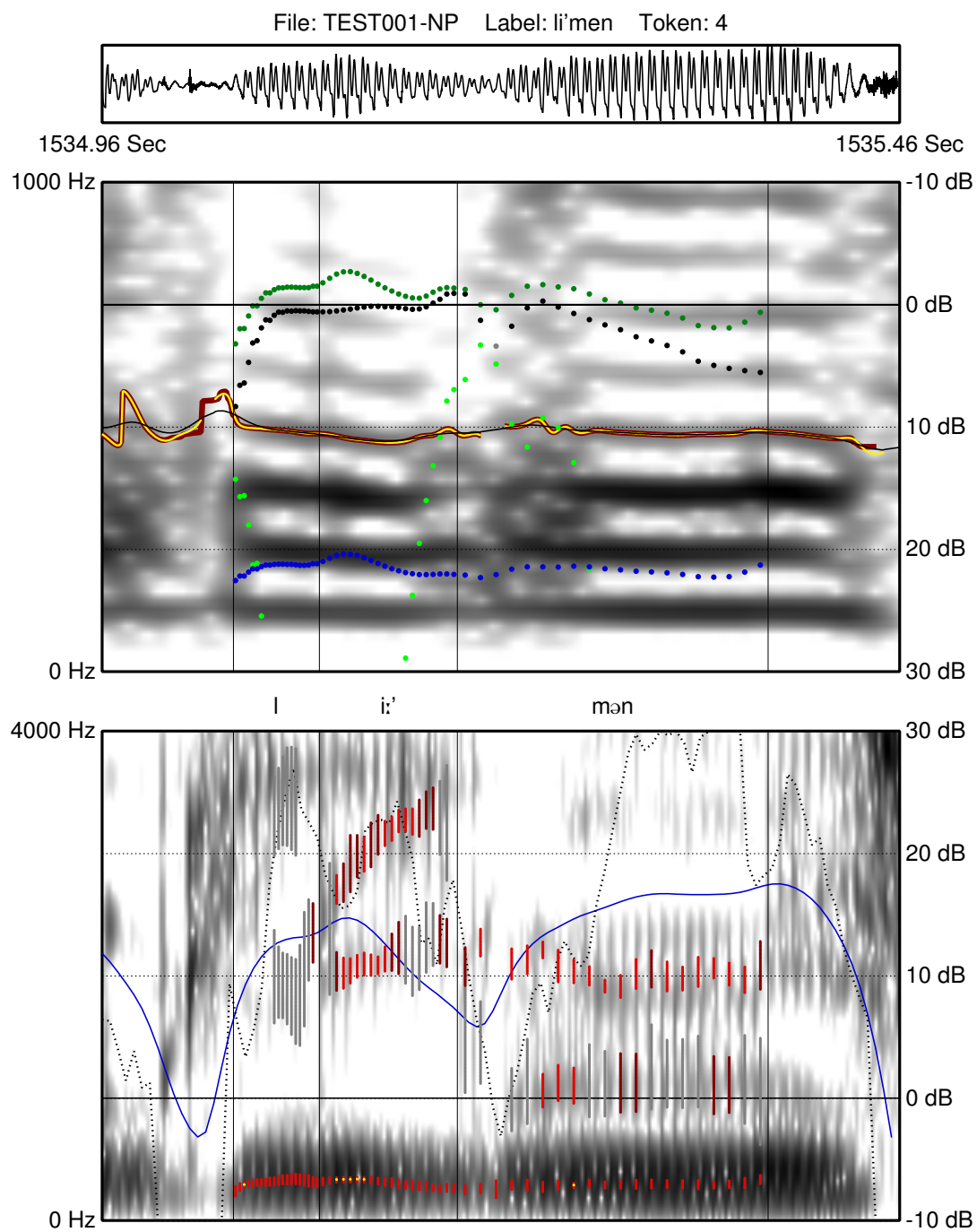


Figure 8.11. limet token nr. 2.



Figur 8.12. li'men token nr. 2.



Figur 8.13. li'men token nr. 4.

8.6 lamm'et vs. lammet

Det sidste ordpar vi vil se på individuelt, er lammet – lamm'et som er et eksempel på stød på konsonant. I figurerne 8.14 og 8.15 er gengivet to eksempler på den stødløse part lammet. Figurerne 8.16 og 8.17 gengiver to eksempler på modparten med stød: lamm'et.

Grundtonen følger i begge ord det ventede lav-høj forløb. I begge ordene har grundtonen sit minimum i slutningen af den betonedede vokal (jeg ser bort fra det lille og meget kortvarige dyk i f_0 ved overgangen mellem [a] og [m]). I ordene uden stød forekommer den højeste f_0 , med lidt variation, i slutningen af den ubetonede stavelse [əð] ved overgangen til det efterfølgende ord. I ordene med stød forekommer den højeste f_0 i begyndelsen af den ubetonede stavelse [əð] stort set sammenfaldende med det tidspunkt hvor svingningerne igen bliver jævnt periodiske. Det tonale forløb i trykgruppens første to stavelser er altså igen en smule mere sammenpresset i tid i ordene med stød. Der er stort set tale om samme niveauer i ordene med som uden stød.

Jeg kan ikke se nogen systematisk forskel mellem ordene med og uden stød hvad angår førsteformantens frekvens, men der er måske en tendens til at førsteformantens båndbredde er mindre i ordene uden stød. (Båndbredden for F1 i [m]⁵⁵ er lav i token 2, 4 og 5 af lammet og kun tilsvarende lav i token 1 af det stødte lamm'et, se appendix H.X).

Der ser ikke ud til at være nogen systematisk forskel med hensyn til varighed for hverken [m] eller [a].

I ordene uden stød (lammet token 2 og 4 i figur 8.14 og 8.15) ses:

Den laveste resonansfrekvens (F1) ligger ved ca. 650 til 750 Hz i [a] og ca. 250-300 Hz i [m]. Dermed ligger F1 i [m] i området hvor H2 og H3 befinder sig. Det er derfor ikke overraskende at BED, H1:H2 og i mindre omfang CoG stiger fra [a] til [m] i takt med at F1 bevæger sig nedad.

H1:H2 udviser et ret svingende forløb i [m] der kunne hænge sammen med om og i hvilket omfang H2 berøres af F1. Det afhænger igen af både F1's båndbredde som ifølge LPC-analysen varierer igennem [m] (se formantplottene i ramme III), og af grundtonen som stiger i løbet af [m]. Men det er også muligt at det er nasalens mere komplekse resonanssystem der spiller ind. På grund af koblingen til næsehulen er nasaler karakteriseret ved et resonanssystem der indebærer både poler (forstærkende) og nuller (dæmpende) som kan ligge ret tæt op ad hinanden. Det laveste nasale pol- og nul-par kan ligge så lavt som 200-400 Hz (afhængig af talerens fysiologi. Se fx Stevens 2000 p. 190, og pp. 494-499 for flere detaljer). Hvis H2 passerer hen over det

⁵⁵ Jeg omtaler her for nemheds skyld den laveste resonansfrekvens i sonoranten [m] som F1 altså som en formant. Er man puristisk indstillet bruges termen formant kun om (de klangmæssigt formende) resonansfrekvenser i mundhulen, og kun når der ikke er nogen akustisk kobling til – og dermed bidrag fra – næsehulen. Se fx Stevens 2000, kapitel 3, note 2.

nedre nasale pol-nul-par som følge af ændringer i grundtonen, så vil styrken af H2 naturligvis variere, og dermed påvirke H1:H2-forholdet. Her er det værd at huske på at LPC-analysen er baseret på en all-pole model og dermed ikke er gearret til at beskrive resonansforholdene i nasaler. Det betyder at de formantværdier der er plottet i nasalen, er at betragte som en grov tilnærmelse der markerer de mest fremtrædende frekvensbånd. Det forslag jeg præsenterede ovenfor om en mulig effekt af F1 på H2, må derfor tages med forbehold.

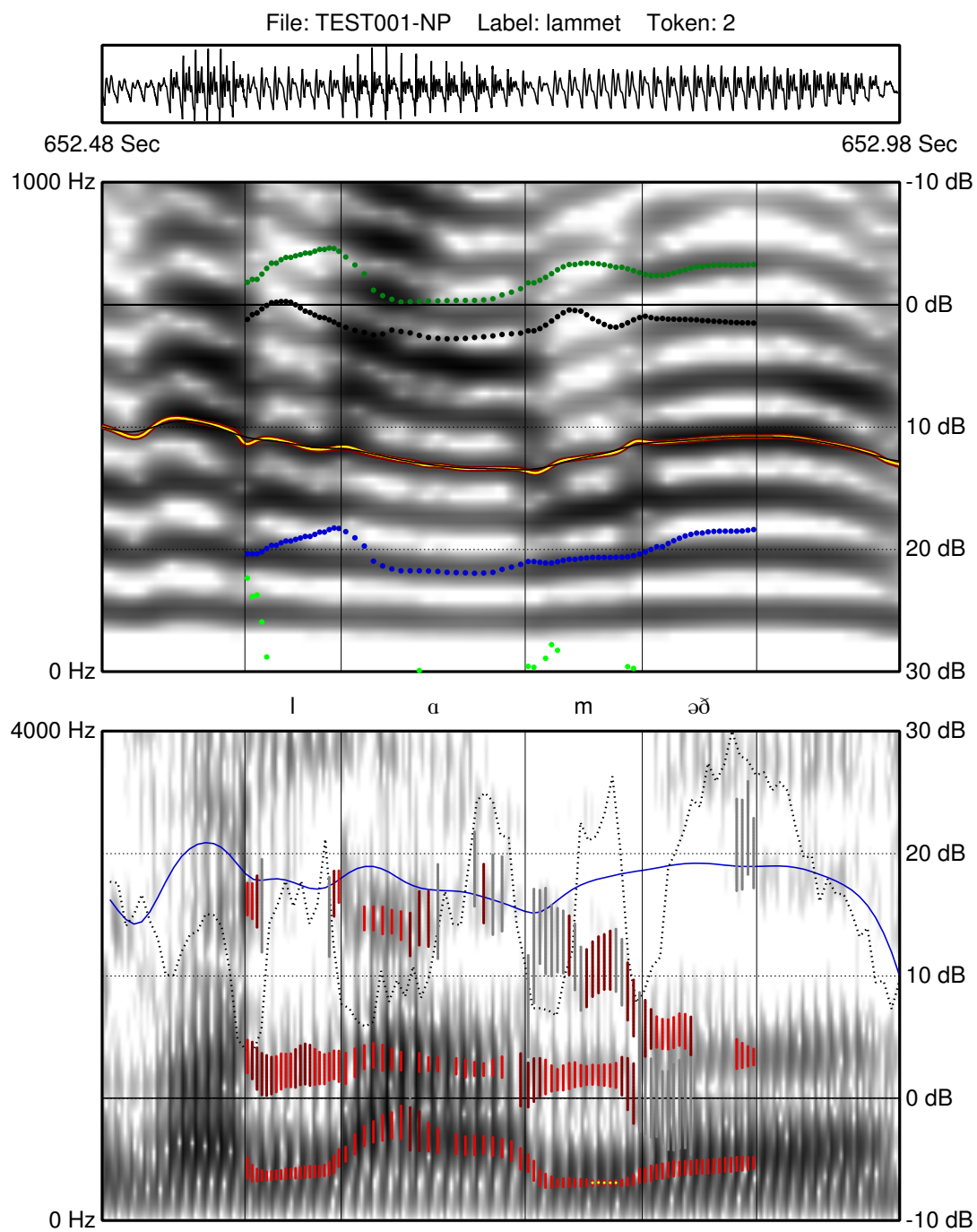
I sammenligning med H1:H2 udviser BED et mere jævnt forløb, men i 2 ud af de 5 tokens ses en markant top i løbet af [m], se fx lammet^{token 2} i figur 8.14. For CoG ses i 4 ud af 5 tilfælde et jævnt eller svagt stigende forløb i løbet af [m]. Kun i lammet^{token 5} udviser CoG-forløbet antydningen af en top midt i [m] (se evt. appendix H.IX). Jeg hælder derfor til den overbevisning at de udsving som H1:H2 og i mindre grad BED udviser i løbet af [m], ikke har med ændringer i stemmekildens kompression at gøre, men i hovedsagen er et produkt af hvordan overtonerne placerer sig i forhold til den nasale sonorants resonanssystems poler og nuller. Ved bedømmelsen af kompressionen i nasalen vil jeg derfor lægge mest vægt på CoG-forløbet.

Til sammenligning ses i modparten med stød følgende:

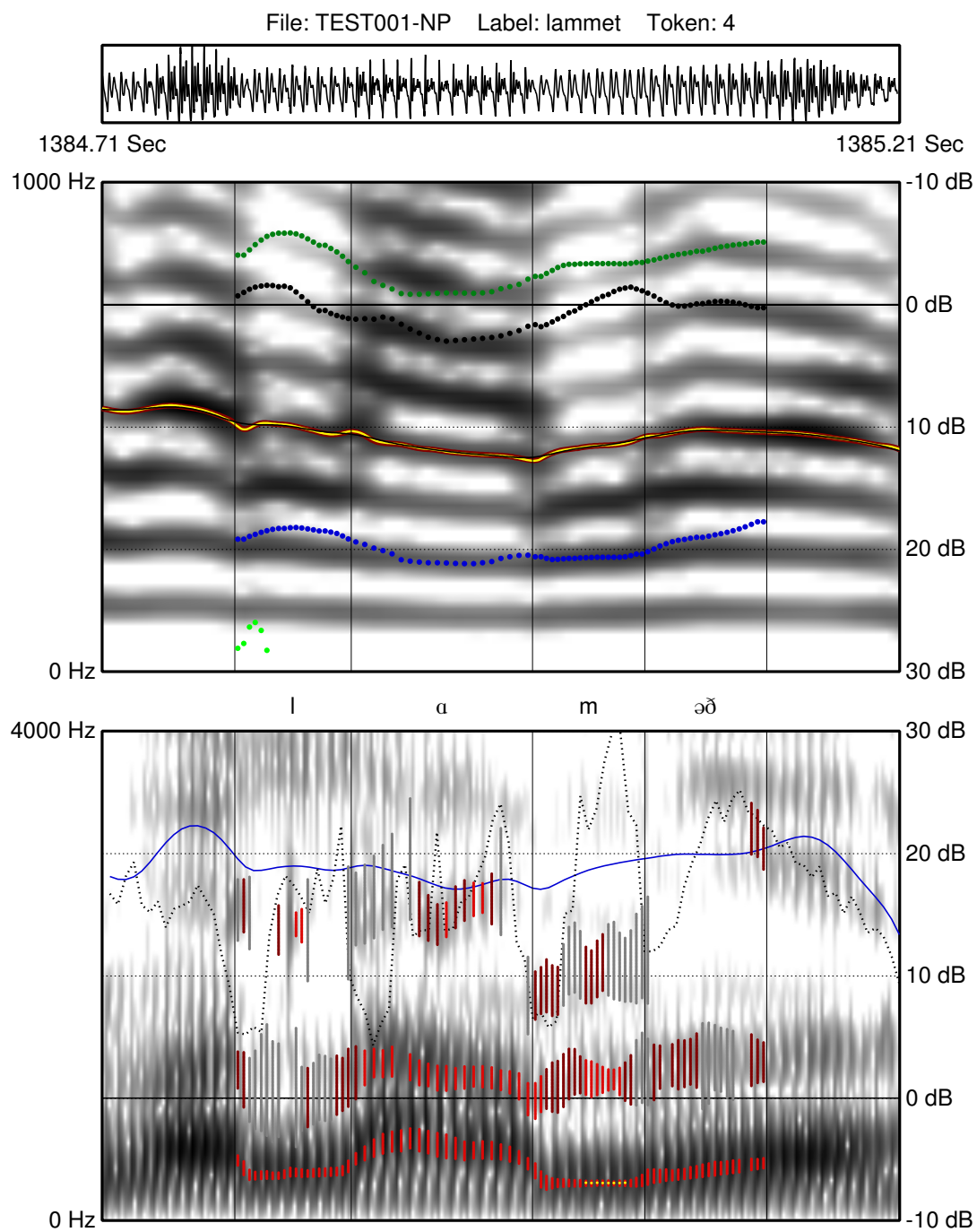
I 4 ud af de 5 tokens med stød stiger kompressionen i løbet af vokalen fra samme niveau (token 2 og token 3, sidstnævnte gengivet i figur 8.16) eller et lavere niveau (token 4 og i særdeleshed token 5, sidstnævnte gengivet i figur 8.17) til et højere niveau (token 2, 3 og 5) sammenlignet med ordene uden stød. Stigningen begynder før det kraftige fald i F1 ved overgangen til [m], så jeg vurderer at der er tale om en genuin ændring i kompressionen. Der ses også stigende kompression i løbet af vokalen i ét af de fem tokens uden stød (token 3, se appendix H.X).

Vurderet ud fra CoG der lader til at være mindst påvirket af F1 (se ovenfor), ser det ud til at kompressionen er kraftigst tæt ved overgangen fra vokal til [m] (token 2, 3 og 4, se token 3 i figur 8.16 som eksempel) umiddelbart før det tidspunkt hvor intensitetskurven har sit minimum. I token 5 (figur 8.17) er kompressionen formodentlig højest i [m] et umiddelbart efter intensitetsdykket, men kompressionen er også høj umiddelbart før dykket. Kompressionen lader således til typisk at være højest umiddelbart før eller i begyndelsen af den stødbærende sonorante konsonant.

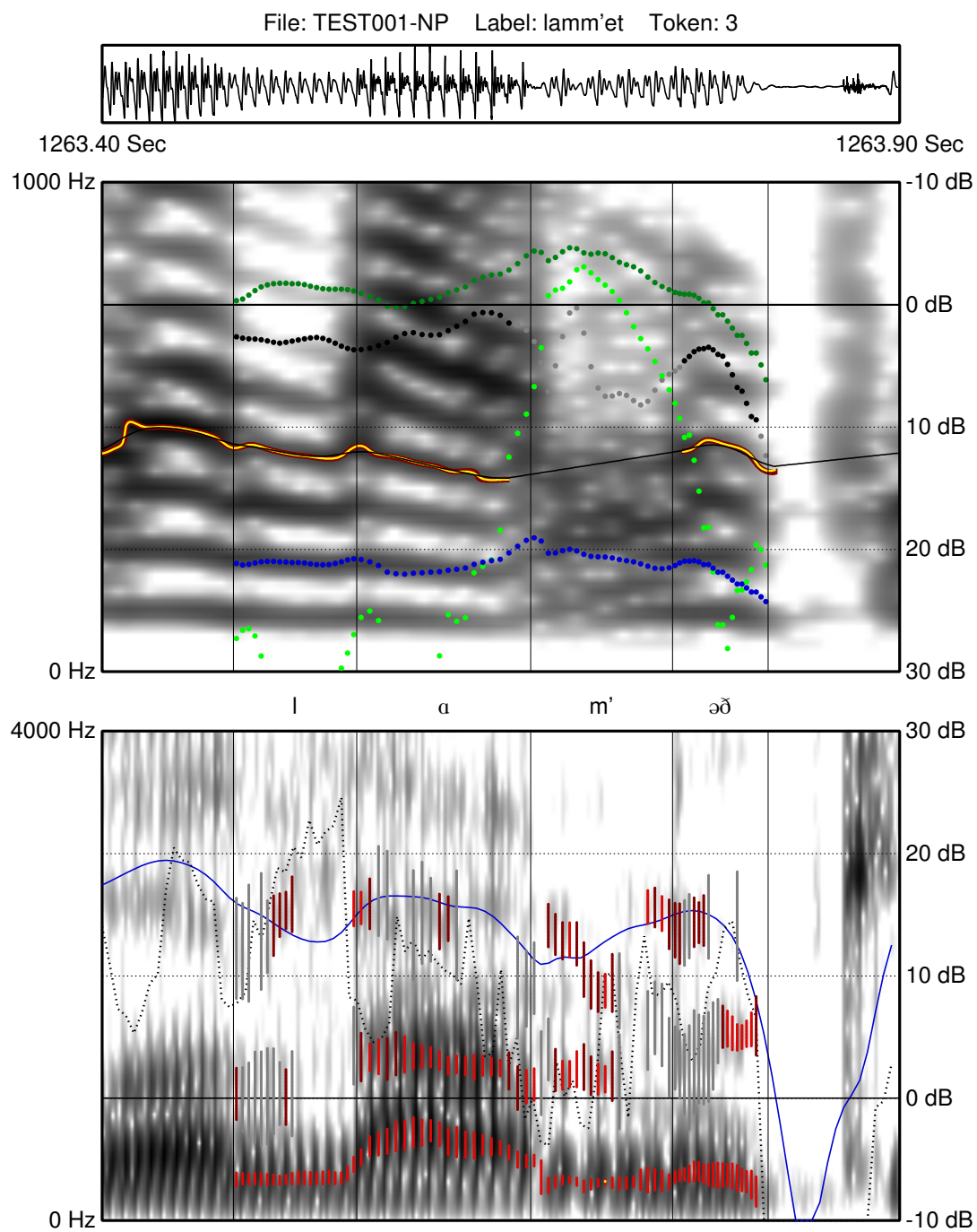
Det er bemærkelsesværdigt at intensitetsforløbet i lamm'et minder meget om intensitetsforløbet i ma'ler ved at det minimum der ses i begge tilfælde, forekommer i begyndelsen af den sonorante konsonant. I begge tilfælde kunne man have ventet at finde intensitetsminimummet tæt ved slutningen af det stødbærende segment hvilket i lamm'et ville betyde nær slutningen af det stødte [m].



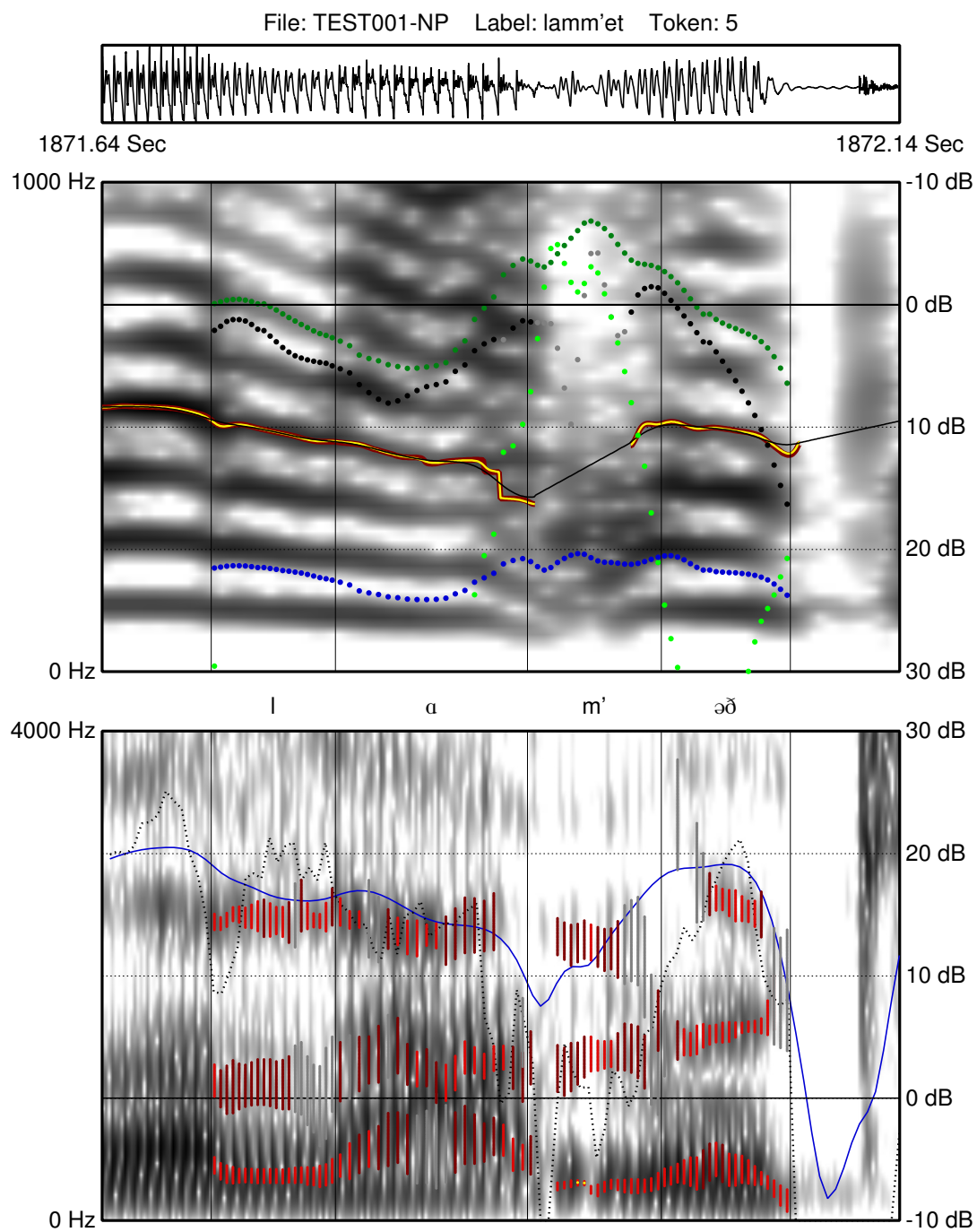
Figur 8.14. lammet token nr. 2.



Figur 8.15. lammet token nr. 4.



Figur 8.16. lamm'et token nr. 3.



Figur 8.17. lamm'et token nr. 5.

8.7 Rekapitulering af hypotesen og de akustiske delhypoteser

Gennemgangen af eksempler har vist at de akustiske markører i mange tilfælde ikke følges ad som man måtte vente ud fra de akustiske delhypoteser. Det forekommer derfor relevant både at tage stilling til om de akustiske analyser som helhed styrker eller svækker hypotesen om stød som en stemmekvalitetsgestus, og at tage stilling til de enkelte akustiske delhypoteser hver for sig.

For overskuelighedens skyld gentages (hoved)hypotesen nedenfor, sådan som den blev præsenteret i kapitel 4, og derefter følger en mere udfoldet beskrivelse af koblingerne mellem hovedhypotesen og de fire akustiske delhypoteser som blev skitseret i slutningen af kapitel 5.

I det følgende vil jeg foreslå en måde at indpasse den labile tilstedeværelse af knirk i en model for stød uden at labiliteten i tilstedeværelsen knyttes til mere eller mindre distinkt udtale. Jeg tager udgangspunkt i en traditionel rangordning af stemmekvaliteter ud fra den glottale åbningsgrad som den præsenteres hos Ladefoged (1971).

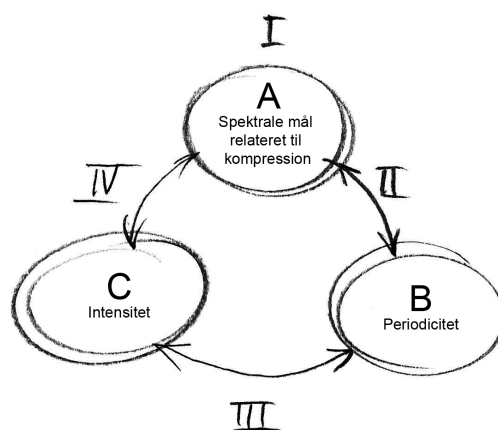
Åbningsgrad	Fonetisk term	Kompression	
lukket	glottal stop		
.	creak		
.	creaky voice		
.	pressed voice	↑	Overkomprimeret Hyperfunktionel ~ lav OQ
.	modal voice	•	Ideal / optimal
.	breathy voice	↓	Underkomprimeret Hypofunktionel ~ høj OQ
.	whisper		
åben	breathing		

Tabel 4.1. Stemmekvaliteter ordnet efter glottal stricture (åbningsgrad). Med udgangspunkt i Ladefoged 1971.

Hypotesen er at stød udtrykkes gennem en relativt kortvarig forandring af stemmekvaliteten i retning af en presset evt. knirket stemmekvalitet og efterfølgende tilbage til mindre presset stemmekvalitet. Stød ansues dermed som en dynamisk stemmekvalitetsgestus. Et velformet stød tænkes at indebære et passende stort udsving i stemmekvaliteten over et passende (kort) tidsrum. Om der forekommer knirk i forbindelse med stød eller ej, tænkes at afhænge af hvor højt oppe på skalaen stødforløbet ligger.

For at kunne teste hypotesen med akustisk fonetiske metoder må den som tidligere nævnt reformuleres i akustisk fonetiske termer. Det fører til fire delhypoteser som ekspliciterer en række akustiske forhold og sammenhænge som man må forvente at kunne iagttage hvis stød udfolder sig som et stemmekvalitetsforløb på den måde det er formuleret i hypotesen. De fire delhypoteser formuleres med reference til tre konkrete (målbare) akustiske forhold:

- A) Mål der relaterer sig til spektral balance som udtryk for kompression
(Her konkret H1:H2, BED og CoG.)
- B) Periodicitet
(Her skelnes mellem regelmæssig og uregelmæssig. Uregelmæssig dækker over diplofoniske og aperiodiske stemmelæbesvingninger.)
- C) Intensitet.



Figur 8.18. Illustration af koblingen mellem de akustiske korrellater for kompression (H1:H2, BED og CoG), periodicitet, intensitet og de fire delhypoteser.

De fire akustiske delhypoteser vedrørende realisationen af stød lyder:

- I. Det forventes at de akustiske mål der relaterer sig til spektral balance som udtryk for kompression, vil udvise et stigende-faldende forløb. I tilfælde hvor stemtheden afbrydes umiddelbart efter stødet forventes et stigende-afbrudt forløb.
- II. Det forventes at de akustiske mål der relaterer sig til spektral balance som udtryk for kompression, vil indikere højere kompression i passager med uregelmæssige svingninger sammenlignet med passager med regelmæssige svingninger.

- III. I ord med stød der realiseres med uregelmæssige svingninger, forventes det at det tidspunkt hvor intensitetsforløbet indikerer at den glottale konstriktion er kraftigtst, falder inden for passagen med uregelmæssige svingninger.
- IV. Det forventes at de akustiske mål der relaterer sig til spektral balance som udtryk for kompression, vil indikere at kompressionen er højest nær det tidspunkt hvor intensitetsforløbet indikerer at den glottale konstriktion er kraftigtst.

8.7.1 Fire spørgsmål: specificering og eksempler

Afprøvningen af hypotesen operationaliseres ved at formulere spørgsmål til hver af de akustiske delhypoteser. I det følgende illustreres det at der for tre af de fire delhypotesers vedkommende både er tilfælde hvor der klart må svares bekræftende på spørgsmålene, og tilfælde hvor der klart må svares afkræftende.

8.8 Delhypotese I

Udviser de akustiske mål der relaterer sig til spektral balance som udtryk for kompression et stigende-faldende forløb? – eller sekundært et stigende-afbrudt forløb i tilfælde hvor stemtheden ikke fortsætter efter stødet?

Spørgsmålet besvares ud fra en samlet vurdering af H1:H2, BED og CoG, men med større vægt på BED og CoG. I behandlingen af dette delspørgsmål ses der bort fra om svingningsmønsteret er regelmæssigt eller ej. Dvs. at jævnt periodiske (modale), diplofoniske og kaotiske passager principielt betragtes på lige fod. Der fokuseres på om der kan ses et overordnet stigende-faldende forløb, og dermed ses der bort fra mindre lokale dyk og uregelmæssigheder. Der ses ikke meget nøje på timingen af forløbet så længe toppen af forløbet forekommer inden for eller i rimlig nærhed af segmentet med stød.

Vi vil først se på nogle eksempler hvor de spektrale mål indikerer stigende-faldende kompression:

ma'ser^{token 2} figur 8.5 og ma'ler^{token 2} figur 8.2 er to eksempler på tokens hvor BED og CoG tydeligt giver en indikation af et stigende-faldende forløb i vokalen med stød. H1:H2 udviser antydningen af et stigende-faldende forløb, klarest i ma'ser^{token 2} figur 8.5, og til dels også i ta'g^{token 5}. Generelt er der meget bedre overensstemmelse mellem H1:H2-forløbet og BED- og CoG-forløbene i tokens uden stød og (dermed) uden uregelmæssige svingninger.

ta'ber^{token 5} i figur 8.10 er et eksempel på stigende-afbrudt forløb, idet stemtheden slutter brat ved overgangen fra den betonedede vokal til lukket i [b].

I de følgende to eksempler indikerer de spektrale mål et forløb i vokalen med stød der ikke kan beskrives som stigende-faldende kompression. Der er altså tale om negativ evidens i forhold til hypotesen.

bar'sel^{token 2} figur 8.18 er et eksempel hvor der ser ud til at være jævnt faldende kompression igennem vokalen. Hele vokalen er præget af uregelmæssige svingninger, så CoG er det mest sikre mål at bedømme ud fra, men det bestyrker vurderingen at BED udviser et tilsvarende faldende forløb.

I a'sen^{token 2} figur 8.19 ses et tydeligt to-toppet forløb. Der er god overensstemmelse mellem forløbet af BED og CoG igennem hele vokalen, og i den del af vokalen hvor der er regelmæssige svingninger, indikerer H1:H2 samme forløb som BED og CoG.

8.9 Delhypotese II

Indikerer de akustiske mål der relaterer sig til spektral balance som udtryk for kompression, at kompressionen er højere i passager med diplofon⁵⁶ eller knirk sammenlignet med passager med regelmæssige svingninger?

Det andet delspørgsmål angår således timingen af den kraftigste kompression i forhold til en eventuel passage med irregulære svingninger. Spørgsmålet besvares ved at sammenholde information om periodiciteten der primært udledes af smalbands-spektrogrammet, med de spektrale mål der kan relateres til kompressionen. I besvarelsen af spørgsmålet må vægten lægges på CoG og BED der ikke stiller de samme krav til stabil periodicitet som H1:H2 gør, og derfor kan give en indikation af om kompressionen stiger eller falder i passager med uregelmæssige svingninger.

Ud over at der kan være forhold der influerer på hvor pålidelige indikatorer for stemmens kompression H1:H2, BED og CoG kan regnes for at være, så er det en forudsætning for at kunne besvare spørgsmålet at der både er modale og ikke-modale svingninger i det relevante tidsrum. Det er ikke muligt at besvare spørgsmålet for tokens med stød uden uregelmæssige svingninger, eller omvendt tokens hvor de uregelmæssige svingninger er udbredt over for lange stræk.

Eventuel negativ evidens vil ikke kun have relevans for afprøvningen af den opstillede model, men vil samtidig udgøre en udfordring for den underliggende model, det vil sige den rangordning af stemtheds måder som Ladefoged præsenterer: Det er et integreret aspekt ved Ladefogeds hierarki at kompressionen er højere i knirk end den er i modal stemme. (Det er i øvrigt den almindeligt gældende opfattelse at knirk er kendetegnet ved høj kompression).

8.9.1 Positive eksempler:

I ta'g^{token 5} figur 8.20 er BED og CoG-værdierne stigende igennem det meste af vokalen og tydeligt højere i vokals sidste del hvor svingningsmønsteret er uregelmæssigt. Tolkningen bestyrkes af at H1:H2 viser den samme opadgående tendens i den del af vokalen hvor svingningerne er regelmæssige nok til at H1:H2 kan beregnes.

Også i ta'ber^{token 5} der blev gengivet i figur 8.10, ses tydeligt højere BED- og CoG-niveauer i den sidste del af vokalen hvor stemmelæbesvingningerne er uregelmæssige. I begge disse tilfælde kan der derfor svares ja til spørgsmålet.

8.9.2 Negative eksempler:

I ma'ler^{token 2} figur 8.2 og ma'ser^{token 2} figur 8.5 indikerer BED- og CoG-forløbet derimod at kompressionen er højere omtrent midt i vokalen hvor stemmelæbesvingningerne er regelmæssige, og lavere i vokals sidste del hvor der ses irregulære svingninger. I disse to tilfælde må der derfor svares nej til spørgsmålet.

⁵⁶ Som nævnt i 4.6 antager jeg at diplofon kan placeres imellem pressed voice og creak (knirk).

8.10 Delhypotese III og IV

Dyk eller fald i intensiteten er et af de mest robuste akustiske træk ved stød. Ifølge Fischer-Jørgensen 1989a sættes dette fald eller dyk i intensiteten i forbindelse med en konstriktion i glottis (der eksplicit ikke betegnes som et glottalt lukke). Kobler man Fischer-Jørgensens observationer med Ladefogeds hierarki, betyder det at intensitetsforløbet må give en indikation af hvor i en stavelse med stød man må forvente at kompressionen er højest: Den mest komprimerede stemmekvalitet i stød må forekomme der hvor intensiteten er lavest. Eller sekundært, i tilfælde af at konstriktionen i stødet er så kraftig at stemmelæbernes svingninger momentant stoppes, må man forvente at finde den mest komprimerede stemmekvalitet i umiddelbar nærhed af det tidspunkt hvor intensiteten er lavest.

Det tredje delspørgsmål angår timingen af intensitetsdykket i ord med stød der realiseres med uregelmæssige svingninger:

Falder det tidspunkt hvor intensitetsforløbet indikerer at den glottale konstriktion er kraftigst inden for en eventuel passage med uregelmæssige svingninger?

I det analyserede materiale ser det ud til at være tilfældet for alle tokens af ord med stød der udviser uregelmæssige svingninger. Se fx figur 8.2 *ma'ser* eller figur 8.16 *lamm'et*.

Det fjerde delspørgsmål angår timingen af den maksimale kompression i forhold til det karakteristiske intensitetsdyk:

Indikerer de akustiske mål der relaterer sig til spektral balance som udtryk for kompression, at kompressionen er højest nær det tidspunkt hvor intensitetsforløbet indikerer at den glottale konstriktion er kraftigst?

8.10.1 Eksempler på tokens med en positiv evaluering:

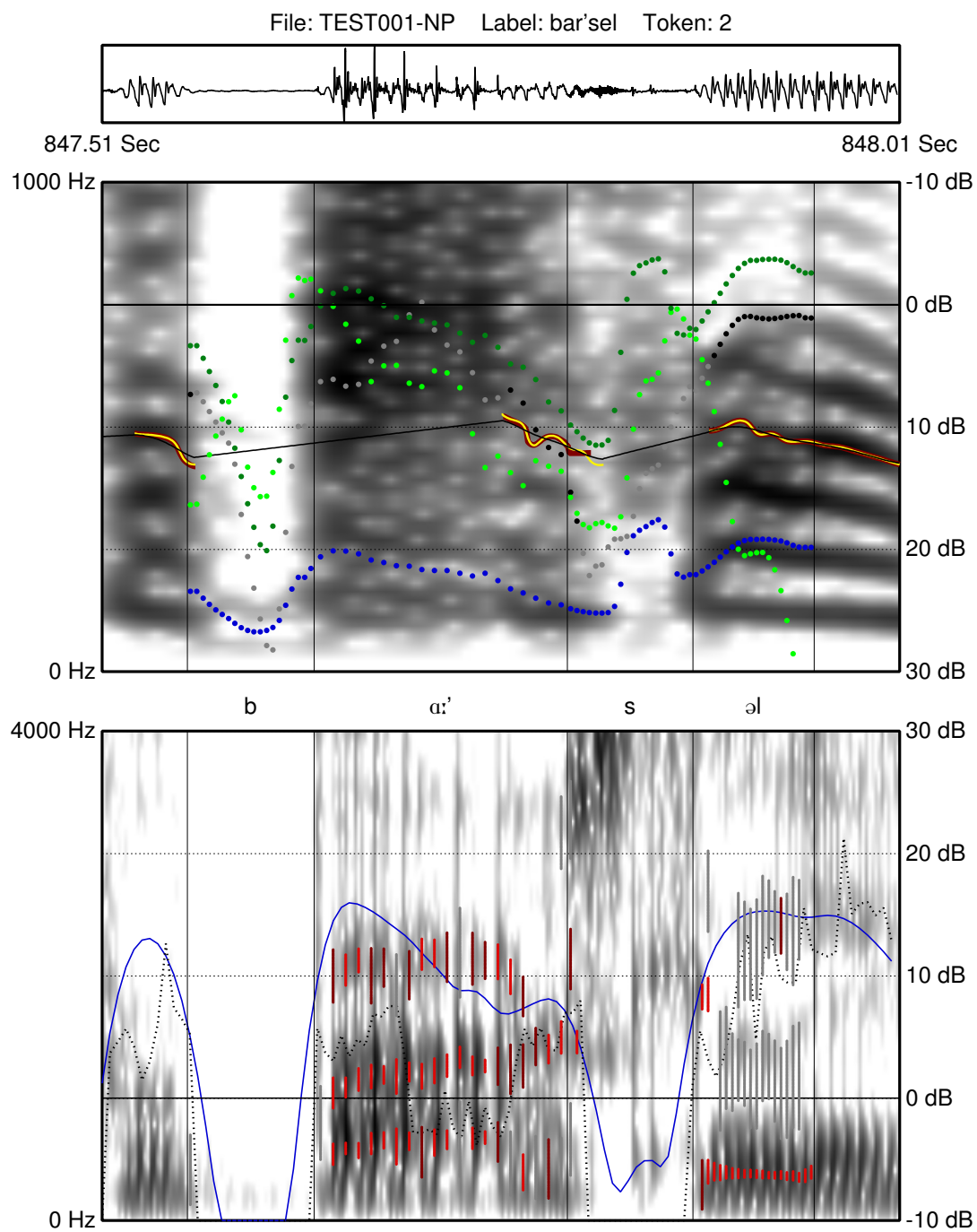
ta'g^{token 5} (figur 8.20) og *ta'ber*^{token 5} (figur 8.10) er begge eksempler på tokens hvor de spektrale mål indikerer at kompressionen er højest nær det sted hvor intensiteten dykker:

Intensitetskurven i *ta'g*^{token 5} (figur 8.20) viser et tvedelt dyk, med to tæt placerede minima. De højeste BED- og CoG-niveauer forekommer mellem de to minima. Det har derfor ikke nogen praktisk betydning for vurderingen af sammenhængen mellem kompression og intensitet hvilket af de to man går ud fra.

I *ta'ber*^{token 5} der blev gengivet i figur 8.10, når BED og CoG de højeste niveauer sidst i vokalen, tæt på det sted hvor intensitetsforløbet går stejlt nedad.

8.10.2 Eksempler der bedømmes negativt i forhold til spørgsmålet:

Negativ evidens er når der er tydelig tidsmæssig adskillelse mellem intensitetsdykket (minimum) og den højeste kompression. De to tokens *ma'ler*^{token 1} (figur 8.3) og *ma'ler*^{token 2} (figur 8.2) er begge eksempler på at intensitetsforløbet og de spektrale mål indikerer at den kraftigste kompression ikke forekommer på samme tidspunkt som den glottale konstriktion: I begge tilfælde forekommer det karakteristiske dyk i intensiteten i den sonore konsonant der efterfølger vokalen, mens de spektrale mål indikerer at den kraftigste kompression ses omtrent midt i vokalen i *ma'ler*^{token 2} (figur 8.2) og endda lidt tidligere i *ma'ler*^{token 1} (figur 8.3).



Figur 8.18. bar'sel token nr. 2.

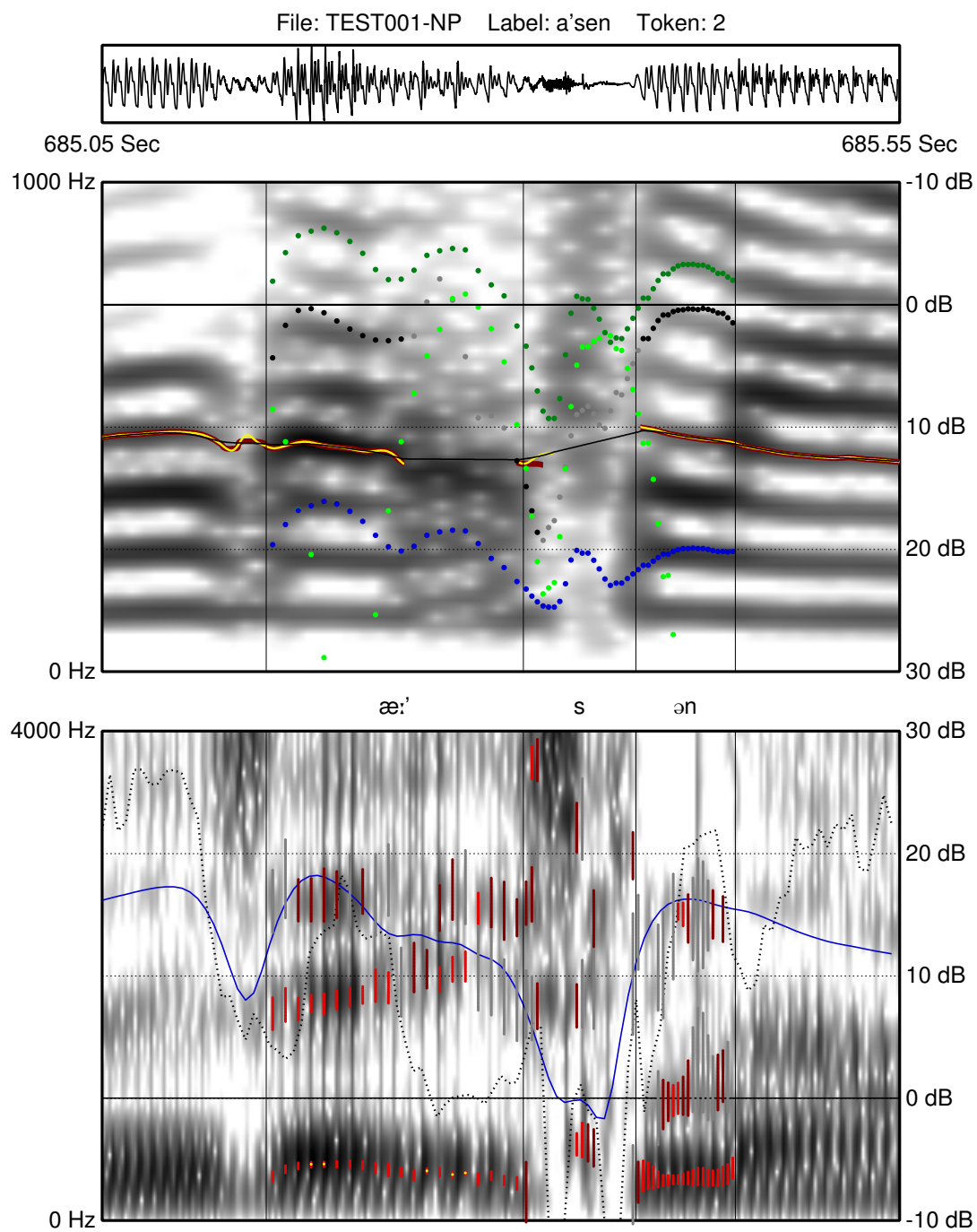


Figure 8.19. a'sen token nr. 2.

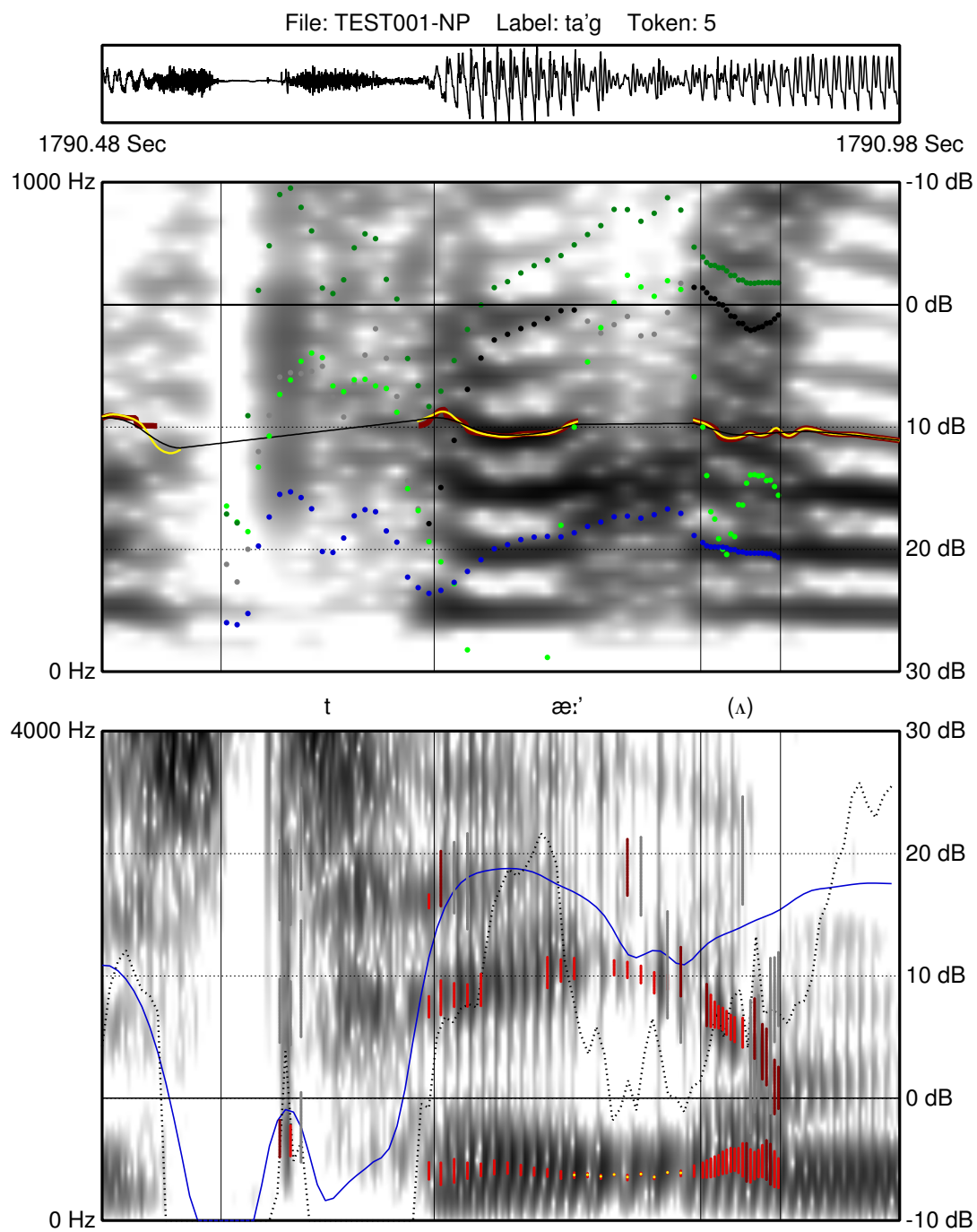


Figure 8.20. ta'g token nr. 5.

8.11 Operationalisering af besvarelse af delspørgsmål

De fire delspørgsmål afklares hver for sig. Hvis der for et givet token kan svares ja til alle fire delspørgsmål, kan det betragtes som positiv evidens for den foreslåede model. Hvis der for et givet token må svares nej til ét eller flere af de fire spørgsmål, betyder det at det må betragtes som tegn på at den foreslåede model ikke holder. Som ved enhver afprøvning af en hypotese er det væsentligste at få identificeret evidens som strider mod hypotesen. Hypotesen som helhed må afvises hvis en substantiel andel⁵⁷ af tokens udviser tegn der må betragtes som negativ evidens.

8.12 Svarkategorier

Svarene på de fire delspørgsmål er inddelt i tre kategorier: 'ja', 'nej' og 'usikker'. Kategoriseringen (besvarelsen af delspørgsmålene) er foretaget ad flere omgange. I vurderingsforløbet er der anvendt et mere fintmasket net med syv svarkategorier som efterfølgende er reduceret til tre. Relationen mellem svarkategorierne kan illustreres således:

- Nej
 - Bestemt nej
 - Mindre sikkert nej
- Ja
 - Bestemt ja
 - Mindre sikkert ja
- Usikker
 - Usikker
 - Temmelig usikker
 - Ubestemmelig

Kategorien 'ubestemmelig' rummer bl.a. tokens som helt sikkert hverken kan klassificeres som 'ja' eller 'nej'.

Som tidligere beskrevet beror besvarelsen af delspørgsmålene på en samlet vurdering af de akustiske målinger. Der har selvsagt været tilfælde hvor det har været vanskeligere at træffe en afgørelse end i andre. Når der har været tvivl, har det i hovedsagen drejet sig om hvorvidt svaret var 'ja' over for 'usikker', eller 'nej' over for 'usikker'. I disse tilfælde er min kategorisering gået i retning af 'usikker'-kategorien.

Eftersom besvarelsen af delspørgsmålene beror på en vurdering – eller tolkning om man vil – af de akustiske målinger, kan der ikke være tale om en objektiv proces. I den bedste af alle verdener burde man derfor få flere uafhængige bedømmere til at forholde sig til målingerne. Men da der ret hurtigt tegnede sig et billede af at

⁵⁷ Principielt skal der selvfølgelig kun et enkelt utvetydigt negativt eksempel til for at afvise hypotesen, men i praksis får en afvisning kun overbevisende tyngde hvis der er passende mange negative eksempler.

materialet ville bestride den opstillede hypotese forekom det tilstrækkeligt at underbygge analysen med illustrationer som efter min vurdering er sikre eksempler på forløb som ikke harmonerer med de formulerede delhypoteser, og dermed ikke harmonerer med hovedhypotesen.

8.13 Selektion

Det materiale der er lavet kurveblade for, består af 308 ord med og uden stød. 210 af ordene hører til minimale og subminimale par med og uden stød. Den resterende del er ord med stød som ikke har en modpart uden stød i materialet.

På baggrund af den indledende gennemgang af kurveforløbene står det klart at der er nogle af single-ordene med stød som det ikke vil give mening at forsøge at analysere på fordi samspillet mellem overtonestrukturen og talekanalens filterfunktion er for kompliceret til at H1:H2-, BED- og CoG-forløbene kan vurderes uden at have en modpart uden stød at holde forløbene op imod. Det drejer sig om et enkelt ord med diftong: fir's og et antal ord med stød på konsonant: æn'der, fil'm, hal', hal'm, mal'm. Omvendt er der et antal single-ord med stød på vokal som godt kan analyseres. Det gælder læ'ner, overma'le, overma'l, far's, fa'gs og fi's.

Det analyserede materiale består dermed af de minimal- og subminimal-par plus single-ord med stød på vokal som fremgår af tabel 8.2 på næste side. De i alt 238 tokens er fordelt på 136 tokens med stød og 102 uden stød.

Oversigt over det analyserede materiale						
Tryk	Segment					
	[ɛæɑ]		[i]		[lm]	
Hovedtryk	lar'men	6	li'men	5	lamm'et	5
	larmet		limet		lammet	
	lar'm	4	smi'ler	5	ball'er	4
	larne		smilet		baller	
	ma'ler	5	smi'l	4		
	maler		smile			
	ma'l	5	fi'n	4		
	male		fine			
	læ'ner	4	ti'	6		
		tie			
	bar'net	5	fi's	4		
	barnlig				
	bar'nsben	5				
	barneben					
	ma'ser	4				
	mases					
	læ'ser	5				
	læser					
	far's	4				
					
	fa'gs	4				
					
	bar'sel	5				
	varsel					
	a'sen	4				
	vasen					
ta'ber	5					
taber						
ta'g	5					
tage						
Bitryk	overma'ler	4			kæmpeball'er	4
	overmaler				kæmpeballer	
	overma'le	7				
					
	overma'l	5				
					
	udta'ler	9				
	udtaler					

Tabel 8.2. Oversigt over analyserede ord. Tallet efter ordet angiver antal tokens. Evt. modpart uden stød er angivet med tekst i gråt. Åben vokal, hovedtryk: lys grøn. Åben vokal, bitryk: mørkere grøn. Snæver vokal: gul. Kort vokal + sonorant: blå.

8.14 Resultater – opdeling af materialet i fire undergrupper

Som det er blevet illustreret i løbet af kapitlet, viser de akustiske analyser i tråd med tidligere undersøgelser at der kan være tydelige forskelle imellem forskellige tokens af samme ord. Der ser også ud til at være variation imellem forskellige ord med stød: Fx ser nogle ord ud til at give mere ensartede tokens end andre. Forskelle på tværs af ord ser ud til at være større eller optræde mere frekvent end forskelle fra token til token af samme ord. Variationen ser ikke ud til at være entydigt bundet til fx segmentel kontekst eller nogen af de øvrige faktorer som materialet dækker. Det kan ikke afvises at der er statistiske tendenser til at bestemte kurveforløb oftere viser sig under bestemte omstændigheder, men det vil kræve et noget større materiale at undersøge om der er sådanne probabilistiske sammenhænge.

En praktisk forskel træder dog tydeligt frem: Som ventet er det vanskeligere at vurdere de akustiske analyser der relateres til kompression i ordene med snæver vokal ([i]) og i ordene med sonorant konsonant på grund af indflydelsen fra den lavestliggende resonansfrekvens⁵⁸. Omvendt er ordene med bitryk lidt lettere at vurdere fordi grundtoneforløbet er mere jævnt i bitrykstavelser.

På vej mod en samlet opgørelse over hvordan materialet stiller sig i forhold til den opstillede hovedhypotese og delhypoteserne, forekommer det derfor rimeligt at opgøre resultaterne for hovedkategorierne hver for sig. Det vil sige at materialet inddeles efter om stødet falder i en stavelse med:

- A. åben vokal (rel høj F1), og hovedtryk
- B. åben vokal (rel høj F1), og bitryk
- C. snæver vokal (lav F1), og hovedtryk
- D. konsonant, hoved- eller bitryk⁵⁹.

I det følgende præsenteres resultaterne for disse fire undergrupper hver for sig.

58 Dvs. F1 eller "F1" i nasaler, jævnfør note 55.

59 Materialet indeholder så få analyserbare tokens med stød på konsonant, at der ikke giver mening at opdele dem yderligere efter trykgrad.

8.15 Resultater for gruppe A

Vi starter med at se på resultaterne for gruppe A, dvs. ord hvor stødet forekommer i en stavelse med en åben vokal og som har hovedtryk. Disse ord er markeret med lys grøn i tabel 8.2. Fordelingen af svar på de tre undersøgte delspørgsmål fremgår af tabel 8.3 nedenfor.

Som tidligere nævnt var det antagelsen at ord med stød på åben vokal ville give de bedste betingelser for at følge akustiske tegn på ændringer i kompressionen. Alligevel må man konstatere at der er en relativt høj andel af ordene i gruppe A hvor et eller flere af delspørgsmålene ikke kan besvares. Det går bedst med hensyn til spørgsmål I hvor andelen er 25 ud af 70, svarende til ca. 36 %, og værst for spørgsmål II hvor andelen der ikke kan besvares, er 36 ud af 70, svarende til ca. 51 % af de undersøgte tokens.

Gruppe A – Åbne vokaler, hovedtryk : [ε æ α]								
n=70	I Op-ned-forløb		II Højere kompres- sion v. uregel- mæss. svingn.		IV Højere kompres- sion op mod intensitetsdyk		Samlet = Hovedhypotese	
Ja	40	57 %	15	21 %	19	27 %	13	19 %
Nej	5	7 %	19	27 %	22	31 %	28	40 %
Ubest.	25	36 %	36	51 %	29	41 %		

Tabel 8.3. Fordelingen af svar på hypotesens delspørgsmål for ord i gruppe A.

Med 5 ud af 70 tokens, svarende til ca. 7 %, er det relativt få tilfælde hvor der må svares nej til spørgsmålet om der ses en op- og ned-adgående ændring i kompressionsniveau, og omvendt kan der svares bekræftende på dette spørgsmål i 40 ud af 70 tilfælde, svarende til ca. 57 %.

For spørgsmål om højere kompression i passager med uregelmæssige svingninger og for spørgsmålet om hvorvidt højeste kompression ses nær intensitetsdykket, gælder det for henholdsvis 19 og 22 ud af 70 tokens at der må svares afvisende på spørgsmålene. Det svarer til negativ evidens i henholdsvis ca. 27 % og ca. 31 % af tilfældene. Der er for begge spørgsmåls vedkommende lidt flere negative end positive svar, men forskellen er kun 3-4 tokens svarende til ca. 4-6 procentpoint.

Samlet set er det 28 ud af 70 tokens, svarende til ca. 40 %, hvor der for ét eller flere af de tre delspørgsmål må svares afvisende, og som derfor må regnes som eksempler der på et eller flere punkter er i modstrid med den foreslåede model. Samtidig er der dog 13 ud af 70 tokens, svarende til ca. 19 %, hvor der kan svares bekræftende på alle tre delspørgsmål, og som derfor ser ud til at følge modellen.

8.16 Resultater for gruppe B

Ordene i gruppe B har bitryk i stedet for hovedtryk og som i gruppe A stød på en vokal med en åben vokalkvalitet. Til forskel fra gruppe A er der kun én vokalkvalitet repræsenteret i gruppe B, nemlig [æ]. Ord i gruppe B er markeret med mørkere grøn i tabel 8.2.

Andelen af tokens i gruppe B der falder i kategorien "ubestemmelig", er mindre for alle tre delspørgsmål sammenlignet med gruppe A, i særdeleshed hvad angår det første spørgsmål. Det er kun for 2 ud af de 25 tokens at kompressionsforløbet ikke kan bedømmes.

Gruppe B – Åben vokal, bitryk : [æ]								
n=25	I Op-ned-forløb		II Højere kompres- sion v. uregel- mæss. svingn.		IV Højere kompres- sion op mod intensitetsdyk		Samlet = Hovedhypotese	
Ja	21	84 %	9	36 %	11	44 %	7	28 %
Nej	2	8 %	5	20 %	6	24 %	10	40 %
Ubest.	2	8 %	11	44 %	8	32 %		

Tabel 8.4. Fordelingen af svar på hypotesens delspørgsmål for ord i gruppe B.

Andelen af tokens hvor der svares negativt på det første spørgsmål, er af samme størrelsesorden som i gruppe A. Andelen af tokens hvor der ses et op- og nedadgående kompressionsforløb, er 21 ud af 25, svarende til omtrent 84 %.

For spørgsmål II og IV er andelen af negative svar måske lidt mindre sammenlignet med fordelingen i gruppe A, mens andelen af positive svar er 15-17 procentpoint højere. Tilsammen betyder det at forholdet mellem negative og positive svar på spørgsmål II og IV i gruppe B ser noget anderledes ud med næsten dobbelt så mange positive som negative svar; henholdsvis 9 over for 5 for spørgsmål II, og 11 over for 6 for spørgsmål IV.

Den samlede andel af tokens der udviser kurveforløb som på et eller flere punkter er i modstrid med hypotesen, er med 10 ud af 25 omtrent samme andel som i gruppe A, nemlig ca. 40 %. Andelen af tokens hvor der kan svares bekræftende på alle tre delspørgsmål, og som derfor ser ud til at følge modellen, er muligvis lidt højere, med 7 ud af 25, svarende til omtrent 28 % (5 tokens ville have svaret til 20 %).

8.17 Resultater for gruppe C

Gruppe C indeholder ord med stød på snæver vokal og dermed lavtliggende F1. De ord der hører til gruppe C, er markeret med gul i tabel 8.2. Små variationer i F1 kan have indflydelse på især H1:H2, men også på BED og CoG, og samtidig kan små ændringer i F1 være svære at se når F1 ligger lavt. Det gælder uanset om man forsøger at aflæse F1 ud fra spektrogrammer, eller om man anvender fx LPC-analyse. Det er derfor forventeligt at der vil være en større andel af tokens hvor der ikke med sikkerhed kan svares på de tre delspørgsmål. For spørgsmål I og II er andelen på henholdsvis 22 og 21 ud af 28 tokens, svarende til omtrent 79 % og 75 %, hvilket som ventet er højere end i gruppe A og B. For spørgsmål IV er andelen der ikke kan besvares, kun lidt højere, med 14 ud af 28 tokens, svarende til 50 %.

Gruppe C – Snæver vokal, hovedtryk : [i]								
n=28	I Op-ned-forløb		II Højere kompres- sion v. uregel- mæss. svingn.		IV Højere kompres- sion op mod intensitetsdyk		Samlet = Hovedhypotese	
Ja	6	21 %	0	0 %	1	4 %	0	0 %
Nej	0	0 %	7	25 %	13	46 %	13	46 %
Ubest.	22	79 %	21	75 %	14	50 %		

Tabel 8.5. Fordelingen af svar på hypotesens delspørgsmål for ord i gruppe C.

Der er ikke nogen negative svar til spørgsmål I hvor andelen i gruppe A og B lå på ca. 8 %. Med et sample på kun 28 tokens kan det være en tilfældighed; to negative svar ville have givet en andel på ca. 8 %.

Der er ikke nogen positive svar til spørgsmål II, og kun ét token vurderes positivt i forhold til spørgsmål IV. Det er væsentligt færre sammenlignet med gruppe A og B. Modsat er andelen der vurderes negativt i forhold til spørgsmål IV med 13 ud af 28, svarende til ca. 46 %, højere end i gruppe A og B.

Den samlede andel af tokens der må regnes som negativ evidens i forhold til hypotesen, er en anelse højere end i de to foregående grupper, med 13 ud af 28 tokens, svarende til 46 %. Andelen der kan regnes som positiv evidens i forhold til hypotesen, er 0 hvilket igen adskiller sig fra resultatet for gruppe A og B hvor andelen lå på 19 % – 28 %.

Samlet set får den større andel af negative svar til spørgsmål IV sammen med de få eller ingen positive svar til spørgsmål II og IV, og dermed fraværet af tokens der er i overensstemmelse med den overordnede hypotese, resultaterne fra gruppe C til at adskille sig noget fra gruppe A og B.

Selvom svarene altså fordeler sig på en lidt anden måde end vi så i gruppe A og B, så er den samlede andel af tokens der giver negativ evidens, dog af samme størrelsesorden som i gruppe A og B.

8.18 Resultater for gruppe D

I gruppe D er ord med stød på konsonant. Ord i denne gruppe er markeret med blå i tabel 8.2. Som nævnt i forbindelse med eksempel-analysen i afsnit 8.6 er resonanssystemet i de sonore konsonanter en tand mere problematisk end snævre vokaler; dels på grund af den lavtliggende første resonansfrekvens "F1", men også fordi der forekommer par af poler og nuller, dvs. resonanser og anti-resonanser der kan ligge ret tæt. Dermed kan selv små ændringer i grundtone (og dermed overtonernes frekvenser) eller omvendt små ændringer i artikulationen (og dermed resonanssystemet) få betydelige konsekvenser for det resulterende spektrum. I sammenligning med gruppe A og B er det derfor ikke overraskende at der i gruppe D er en højere andel af tokens hvor det ikke har været muligt at besvare de tre delspørgsmål. Da behovet for at have matchende par med og uden stød af samme årsager viste sig at være mere uomgængeligt for ord med stød på konsonant, betyder det at opgørelserne i gruppe D er baseret på forholdsvis få tokens. Resultaterne må derfor tages med ydeligere forbehold.

Gruppe D – Sonorant konsonant, hovedtryk eller bitryk : [l m]								
n=13	I Op-ned-forløb		II Højere kompres- sion v. uregel- mæss. svingn.		IV Højere kompres- sion op mod intensitetsdyk		Samlet = Hovedhypotese	
Ja	5	38 %	2	15 %	6	46 %	2	15 %
Nej	1	8 %	0	0 %	1	8 %	1	8 %
Ubest.	7	54 %	11	85 %	6	46 %		

Tabel 8.6. Fordelingen af svar på hypotesens delspørgsmål for ord i gruppe D.

Fordelingen af svar i gruppe D adskiller sig fra de foregående ved at der ikke er nogen negative svar til spørgsmål II og kun et enkelt negativt svar til spørgsmål IV. Omvendt er der en stor andel af positive svar til spørgsmål IV – i samme størrelsesorden som det der blev iagttaget i gruppe B

Hvad angår spørgsmål II og IV, er fordelingen nærmest komplementær til den vi så for [i] i gruppe C. Fordelingen minder mest om den fordeling der viste sig for gruppe B, men dog med en endnu større andel af positive svar til spørgsmål II og IV.

Der er kun et enkelt ud af de 13 tokens som kan siges at bestride hypotesen, mens der er to som er i overensstemmelse med modellen. Med så få tokens i gruppen kan det være et udslag af tilfældigheder at det er den eneste gruppe hvor andelen af tokens der bestrider hypotesen, er under 40 % og samtidig den eneste gruppe hvor der er flere tokens der kan regnes som positiv evidens, end der kan regnes som negativ evidens.

9 Akustisk analyse: resultater

9.1 Evaluering af hovedhypotese

I tabel 9.1 nedenfor er hovedresultatet for de fire grupper samlet sådan at man kan se fordelingen af tokens der udviser et eller flere tegn som er i modstrid med den forelåede hypotese.

Fordeling af negative svar ifht. hovedhypotesen			
Grp.	n	Negativ evidens - antal	Negativ evidens %
A	70	28	40 %
B	25	10	40 %
C	28	13	46 %
D	13	1	8 %

Tabel 9.1. Antal og andel af tokens der udviser tegn der er i modstrid med hovedhypotesen fordelt på de fire hovedgrupper.

For stød på vokal, dvs. grupperne A, B og C, er det 40-46 % af tokens der udviser tegn der ikke er forenelige med hypotesen. For grupperne A, B og C som helhed gælder det 51 ud af 123 undersøgte vokaler med stød, svarende til ca. 41 %.

Med hensyn til stød på konsonant er det kun et enkelt token ud af 13 som udviser tegn der er i modstrid med hovedhypotesen. Det er på nuværende tidspunkt ikke klart hvorfor resultatet i gruppe D skiller sig ud fra de øvrige grupper. Det må dog bemærkes at der netop i gruppe D har været en større andel af tokens hvor det ikke har været muligt at besvare et eller flere af de tre delspørgsmål. (Det samme gælder gruppe C, men altså ikke med samme konsekvens.) Jeg vender tilbage til fordelingen af svar og forskellene imellem grupperne i afsnit 9.5.

Uanset om forholdene rent faktisk er anderledes for konsonanternes vedkommende, og uanset at der er tokens der er vurderet til at være i overensstemmelse med hypotesen, så er andelen af vokaler med stød der er vurderet til at være i modstrid med hypotesen på et eller flere punkter, så høj at det næppe vil være frugtbart at gå videre med en afprøvning af hypotesen på et større materiale. Hypotesen om stød som en stemmekvalitetsgestus (som gentaget i afsnit 8.7) må betragtes som afvist.

9.2 Evaluering af delhypotese II

Fordeling af svar ifht. delhypotese II					
Grp.	n	Positiv evidens - antal	Positiv evidens %	Negativ evidens - antal	Negativ evidens %
A	70	15	21 %	19	27 %
B	25	9	36 %	5	20 %
C	28	0	0 %	7	25 %
D	13	2	15 %	0	0 %
Sum	136	26	19 %	31	23 %

Tabel 9.2. Resultater for delhypotese II fordelt på de fire hovedgrupper.

I ca. 19 % af de undersøgte tokens indikerer de akustiske korrelater for kompression at kompressionen når sit maksimum i løbet af passagen med uregelmæssige svingninger. Men i lidt flere tilfælde (23 %) topper kompressionen før der opstår uregelmæssigheder i stemmelæbesvingningerne. Den højeste kompression følges altså ikke nødvendigvis af uregelmæssigheder i stemmelæbesvingningerne (aperiodicitet eller diplofoni).

Dermed kan den traditionelle forklaring om at uregelmæssigheden i stemmelæbesvingningerne er en konsekvens af høj kompression (se fx Stevens 2000), i hvert fald ikke være enerådende; der må være andre forhold der kan medføre uregelmæssige stemmelæbesvingninger.

Som tidligere nævnt beskriver Slifka 2006 også uregelmæssige stemmelæbesvingninger som forekommer uden at det har med høj kompression at gøre. Der er tale om uregelmæssige svingninger som samtidig er karakteriseret ved relativt højt (frit) glottalt flow.

I forbindelse med stød taler Fischer-Jørgensen 1989a entydigt om tegn på glottal konstriktion. Det sker bl.a. på baggrund af målinger der viser et reduceret glottalt flow (målt ved hjælp af Rothenberg-maske), men også understøttet af fx intensitetsforløbet i det akustiske signal. De uregelmæssige stemmelæbesvingninger som ses i dette materiale og som dårligt kan skyldes høj kompression, kan derfor næppe heller side-stilles med den type uregelmæssige stemmelæbesvingninger der beskrives i Slifka 2006.

Det kunne derfor enten tyde på at der findes endnu en type uregelmæssige svingninger, eller at uregelmæssighed i svingningsmønstret kan forekomme uafhængigt af stemmekildens kompression. Det sidste må efterhånden forekomme mere sandsynligt.

9.3 Evaluering af delhypotese IV

Følger man Fischer-Jørgsensens beskrivelse, kan intensitetsforløbet bruges som indikator for hvornår den glottale konstriktion når sit maksimum. Kobles det forhold med Ladefogeds hierarki, betyder det at man må forvente at finde den kraftigste kompression nær intensitetsdykkets minimum. Men resultaterne for delhypotese IV viser at der ikke nødvendigvis er den ventede sammenhæng mellem intensitetsdykket og kompressionen.

Fordeling af svar ifht. delhypotese IV					
Grp.	n	Positiv evidens - antal	Positiv evidens %	Negativ evidens - antal	Negativ evidens %
A	70	19	27 %	22	31 %
B	25	11	44 %	6	24 %
C	28	1	4 %	13	46 %
D	13	6	46 %	1	8 %
Sum	136	37	27 %	42	31 %

Tabel 9.3. Resultater for delhypotese IV fordelt på de fire hovedgrupper.

I ca. 27 % af de undersøgte tokens er der en god overensstemmelse imellem timingen af den kraftigste kompression og intensitetsdykket. Andelen der vurderes positivt i gruppe C (dvs. i ord med stød på snæver vokal), er dog noget mindre.

I ca. 31 % af de undersøgte tokens er der imidlertid en klar adskillelse i tid imellem de to hændelser, sådan at kompressionen typisk topper omkring 80-100 ms før intensitetsdykkets minimum. Disse forløb harmonerer ikke med delhypotesen.

Eftersom den kraftigste kompression og intensitetsdykkets minimum kan times i forhold til hinanden på (mindst) to forskellige måder, så kan høj kompression og glottal konstriktion ikke være fast koblet med hinanden.

9.3.1 Opsamling på resultaterne for delhypotese II og IV

Resultaterne for delhypotese II og IV fører ikke bare til afvisningen af hovedhypotesen; de tokens som er i modstrid med enten den ene eller begge af delhypoteserne II og IV, kan dårligt passes ind i Ladefogeds hierarki. Dermed udgør de en udfordring for hierarkiets værdi som (universel) systematik over funktionel brug af stemmekvaliteter. Det er et resultat som peger på at der fortsat er behov for videre studier af stemmekildens funktionsmuligheder.

9.4 Observation af variation – to yderpunkter

Svarene til delspørgsmålene tjener ikke kun til at afvise delhypotese II og IV (og dermed hovedhypotesen): Som man kan se af tabel 9.2 og 9.3, så er der for både delhypotese II og IV samlet set en næsten lige stor andel af positive og negative resultater. Man vil også bemærke at der er visse lighedstræk i fordelingen af svar for de to delhypoteser. Især træder den lave andel af positiv evidens for gruppe C og den lave andel af negativ evidens for gruppe D frem i begge tabeller.

I det gennemgåede materiale indikerer timingen af intensitetsdykket at den glottale konstriktion i alle tilfælde falder inden for intervallet med uregelmæssige stemmelæbesvingninger. Dermed bestyrkes delhypotese III. På den baggrund må man formode at der vil være en forholdsvis pæn overensstemmelse imellem fordelingen af positive og negative svar til delspørgsmål II og IV: Hvis kompressionsmaksimummet falder tæt på intensitetsminimummet, så vil kompressionsmaksimummet også (jfr. ovenstående) falde inden for intervallet med uregelmæssige stemmelæbesvingninger. Omvendt, hvis kompressionsmaksimummet forekommer før passagen med uregelmæssige svingninger, så forekommer det også før intensitetsdykkets minimum.

Der behøver dog ikke nødvendigvis at være fuld overensstemmelse mellem svarene til delhypotese II og IV. Hvis der er et langt stræk med uregelmæssige svingninger, kan kompressionsmaksimummet godt forekomme i passagen med uregelmæssige svingninger uden at falde tæt ved intensitetsdykkets minimum. I de tilfælde hvor et af de to delspørgsmål (II eller IV) ikke kan besvares, kan man selvfølgelig ikke sige om der er overensstemmelse imellem dem.

Hvis man opstiller svarene til delhypotese II og IV i en matrice (tabel 9.4 nedenfor), kan man se at sammenhængen mellem svarfordelingerne er i god overensstemmelse med ovenstående beskrivelse.

n=136		Delhypotese IV		
		Ja	Ubest.	Nej
Delhypotese II	Ja	23 17 %	2 1.5 %	1 0.7 %
	Ubest.	14 10 %	49 36 %	16 12 %
	Nej	0 0 %	6 4.4 %	25 18 %

Tabel 9.4. Fordelingen af resultater for delhypotese II og IV.

Der ses ikke nogen tilfælde hvor den kraftigste kompression observeres tæt ved intensitetsdykkets minimum (dvs. *ja* til delspørgsmål IV) samtidig med at den kraftigste kompression forekommer før passagen med uregelmæssige stemmelæbesvingninger (dvs. *nej* til delspørgsmål II). Der er et enkelt tilfælde hvor kompressionsmaksimummet forekommer i passagen med uregelmæssige svingninger (dvs. *ja* til delspørgsmål II), men ikke sammenfaldende med tidspunktet for intensitetsdykkets minimum. Det pågældende token er netop et tilfælde hvor hele vokalen er præget af uregelmæssige svingninger.

Eftersom der stort set ikke er nogen tilfælde hvor der gives modsatte svar på de to delspørgsmål, forekommer det rimeligt at gruppere de tokens der kun er vurderet som ubestemmelig i forhold til ét af de to delspørgsmål, ud fra det delspørgsmål der kunne besvares. Grupperingen er markeret i tabel 9.4 med lys blå og grå. Ud fra fordelingen af svar på delspørgsmål II og IV tegner der sig dermed et billede af to yderpunkter som man tentativt kunne benævne som type P og type N.

De to typer kan beskrives således:

Type P: Kompressionen når sit maksimum meget tæt på det tidspunkt hvor intensitetskurven indikerer at den glottale konstriktion er kraftigst. I det undersøgte materiale betyder det samtidig at den højeste kompression iagttages i en passage hvor stemmelæbesvingningerne er uregelmæssige.

Type N: Kompressionen når sit maksimum før stemmelæbesvingningerne bliver uregelmæssige og (ca. 80-100 ms) før intensitetskurven indikerer at den glottale konstriktion er kraftigst.

Positivt svar til ét eller begge de to delspørgsmål tages som indikation for P-typen (markeret med lys blå i tabel 9.4) og negativt svar til ét eller begge de to spørgsmål tages som indikation for N-typen (markeret med lys grå i tabel 9.4). Tokens der enten er vurderet som ubestemmelige med hensyn til begge delspørgsmål eller hvor svarene er modstridende, regnes som ubestemmelige.

Resultaterne peger således på to forskellige kombinationer af forløb. De tilgængelige data giver dog ikke noget klart vidnesbyrd om hvorvidt der er tale om om glidende variation mellem to yderpunkter eller om der eventuelt er tale om to distinkt forskellige måder at realisere stød på.

9.5 Observation: tegn på at variationen ikke er tilfældigt fordelt

Tabel 9.5 nedenfor viser hvordan de to foreslåede typer fordeler sig over de hovedgrupper det undersøgte materiale er opdelt efter.

Fordeling af type P og N ud fra svar på delspørgsmål II og IV					
Grp.	n	Type P - antal	Type P %	Type N - antal	Type N %
A	70	20	29 %	26	37 %
B	25	12	48 %	7	28 %
C	28	1	4 %	13	46 %
D	13	6	46 %	1	8 %
Sum	136	39	29 %	47	35 %

Tabel 9.5. Fordeling af type N og P over de fire hovedgrupper.

Med 39 forekomster af type P, svarende til ca. 29 %, og 47 forekomster af type N, svarende til ca. 35 %, er de to typer omtrent lige hyppige i materialet som helhed. 50 af de 136 tokens svarende til ca. 36 % kan ikke kategoriseres. Forekomsterne er ikke jævnt fordelt: Type P er hyppigst i forbindelse med bitryk og konsonant og mindst hyppig i forbindelse med stød på snæver vokal. Type N er hyppigst i forbindelse med snæver vokal (med hovedtryk) eller åben vokal med hovedtryk og mindst hyppig i forbindelse stød på konsonant.

De to grupper der skiller sig tydeligst ud, er således gruppe C og D som er ord med stød på henholdsvis snæver vokal og konsonant. Det er kedeligt at de to grupper hvor fordelingen afviger mest, samtidig er de grupper som har været vanskeligst at vurdere. Problemt består i at man ikke kan vide om den andel det har været muligt at evaluere er repræsentativ for gruppen som helhed. I de tilfælde hvor andelen af reelle svar er lille, kan man sige at svarene kun siger noget om "toppen af isbjerget". Problemet er at man ikke kan vide om det er den samme del af "isbjerget" der vender opad i de forskellige grupper. Med andre ord kan man ikke afvise at de forskelle der ses, kan skyldes at vi ser forskellige dele af ellers sammenlignelige "isbjerger".

Det problem løses ikke ved at se på et større materiale; det er nødvendigt at øge svarprocenten. Hvis man vil undersøge fordelingen af de to typer nærmere er det derfor et påtrængende spørgsmål om det kan lade sig gøre at forbedre metoden så den bedre kan håndtere sonorante konsonanter og snævre vokaler.

9.5.1 Opsamling på observationerne vedrørende variation i realisationen

De akustiske analyser der er foretaget, viser en type variation i måden stød kan realiseres på, som ikke tidligere er beskrevet. Det foreslås at variationen spænder over to (yder)typer der tentativt benævnes type P og type N. Yderligere er der visse tegn på at variationen ikke er fuldstændig tilfældig, men at der kan være distributionelle mønstre.

At stød kan have flere fremtrædelsesformer som har et distributionelt mønster, kendes fra beskrivelser af andre danske lekter (aka varieteter), fx beskrivelsen af østfynsk i Andersen 1958 hvor stød ses alternativt at kunne realiseres som/med en stigende tonal bevægelse, men forbeholdt ord med snæver vokal. Senest har Bodil Kyst (2008) peget på at stød i Århusiansk kan realiseres på to måder, der beskrives som tonalt forskellige, og sandsynliggjort at forskellen har betydningsadskillende potentiale.

Hvis de to foreslåede typer P og N kan genfindes hos et større antal talere, og hvis der kan findes statistisk holdbare tegn på systematik i distributionen, så vil der være tale om noget nyt for den variant af dansk som er undersøgt her, dvs. københavnsk rigsmål.

9.6 Observation: diplofoni ser ud til at være udbredt

I det undersøgte materiale har det – lidt overraskende – vist sig at i stort set alle tokens med stød hvor der forekommer uregelmæssige stemmelæbesvingninger, kan en større eller mindre del af den uregelmæssige passage betegnes som diplofonisk.

Også andre har bemærket forekomsten af diplofoniske svingninger i forbindelse med stød. Fx nævner Svend Smith (1944) at der forekommer "regelbunden Uregelmæssighed" (dvs. diplofoni) i nogle ord med stød i forbindelse med "en gradvis Overgang fra de stationære [dvs. regelmæssige] Svingninger til de ikke-stationære" (p. 95).

I de fleste tilfælde betegnes aperiodiske og diplofoniske stemmelæbesvingninger dog under ét som uregelmæssige⁶⁰, og man kan derfor let få den opfattelse at diplofoni skulle være en sjældent forekommende kuriøsitet. Det kan nok til dels tilskrives det forhold at der har været tradition for at vurdere periodicitet ud fra en form for oscillografisk visning af lydsignalet. Periodiciteten kan ofte ikke bedømmes pålideligt ud fra et oscillogram: Svagere diplofoni vil let blive overset, og det er ikke altid muligt at skelne sikkert mellem diplofoni og aperiodicitet. Når periodiciteten vurderes ud fra en spektral analyse der kan vise deltonestrukturen (dvs. smalbandsspektrum), så viser det sig at diplofoniske svingninger optræder ganske hyppigt. Når man tidligere oftest har brugt oscillografiske metoder, så skyldes det sandsynligvis at spektrale analyser historisk set har været meget "dyre" at foretage i sammenligning med oscillogrammer, og man har formodentlig også ment at oscillogrammet gav tilstrækkelig information om talesignalet periodicitet.

En anden faktor der kan have påvirket interessen for at skelne mellem diplofoni og aperiodicitet, er at det fra en auditiv synsvinkel er uklart om det er væsentligt at skelne mellem diplofoni og aperiodiske stemmelæbesvingninger – ikke kun i forbindelse med stød. Det er forholdsvis sparsomt med viden om den auditive forskel på stemme med diplofoniske og aperiodiske svingninger, men en undersøgelse af Gerrat og Krizan (2001) viser at lyttere kan skelne kategorisk mellem stemmer der ud fra akustiske kriterier er vurderet til enten at være præget af diplofoni eller af aperiodicitet. Spørgsmålet om hvilken auditiv status forekomsten af diplofoni måtte have i forbindelse med stød, forekommer derfor relevant.

Akustisk set er forskellen imellem diplofoni og aperiodicitet ubestridelig, og den kommer som nævnt tydeligt til udtryk i forbindelse med spektrale analyser der kan vise deltonestrukturen. Yderligere peger eksperimenterne i afsnit 6.20 på at det kan være væsentligt at skelne mellem diplofoni der indebærer variation i periodetiden (dvs. frekvensmodulation), og diplofoni der kun rammer amplituden (dvs. amplitudemodulation), eftersom forskellen (naturligt nok) afspejles i den spektrale energifordeling.

⁶⁰ I de refererede værker af Fischer-Jørgensen og Grønnum & Basbøll betegnes aperiodiske og diplofoniske svingninger generelt under ét (men med lidt varierende betegnelser, jævnfør gennemgangen i kapitel 3).

10 Afrunding

Når passagen med uregelmæssige svingninger i mange tilfælde ikke udgør toppen af et stigende-faldende kompressionsforløb, så falder ideen om et kompressionsforløb der forener ændringer i pulsform (H1:H2), periodicitet (periodisk-diplofonisk-aperiodisk) og intensitet (=glottal konstriktion) til jorden, og ideen om at det skulle være dette (kompressions-)forløb der skulle udgøre det primære signal om stød, modsiges. Når kompression ikke nødvendigvis hænger sammen med uregelmæssige svingninger og med intensitetsdykket sådan som hypotetiseret, så bidrager kompressionsforløbet ikke til at forklare den variabilitet der ses især med hensyn til forekomsten af uregelmæssige svingninger i forbindelse med stød.

Det oprindelige problem står dermed stadig ubesvaret: hvordan forklares det tilsyneladende misforhold mellem den perceptuelle konstans og den akustiske variabilitet? Hvordan kan noget der "ser så forskelligt ud" lyde som "det samme"?

Og status quo er stadig at der ikke eksisterer en velafprøvet model for – eller karakteristik af – hvad stød er fonetisk set.

Det er dog samtidig tydeligt at der er ny viden at høste ved at se på ændringer i kompressionen i forbindelse med stød, og at muligheden for at følge kompressionsforløbene kan give ny indsigt. Sammenholder man data fra tokens med og uden stød, ser det ud til at de adskiller sig både hvad angår kompressionens forløb og niveau: Kompressionen ændrer sig løbende, men i ord med stød er forløbet mere dynamisk, og kompressionen er generelt højere sammenlignet med ord uden stød. Overordnet set lader det til at ord med stød typisk er karakteriseret ved et kompressionsforløb der kan betegnes som stigende-faldende. Men timingen ser ud til at variere.

Beskrivelsen af kompressionens forløb bibringer dermed information om endnu en form for variation i måden stød realiseres på; enten forekommer den kraftigste kompression og intensitetsdykket meget tæt på hinanden og samtidig med de uregelmæssige stemmelæbesvingninger eller også forekommer den kraftigste kompression noget før intensitetsdykket og før der ses uregelmæssige stemmelæbesvingninger. Tentativt benævnes de to måder: type P og type N.

De indsamlede data giver indikationer af at der kunne være en vis systematik i distributionen af de to foreslåede typer. Det vil være interessant at undersøge nærmere i et større materiale.

11 Referenceliste

- Abrahams, H. (1943). Nogle fonetiske iagttagelser vedrørende Stødet i moderne Dansk.
In Memoriam Kr. Sandfeld (Red. Brøndal, R., Brøndal, V., Møller, C., Olsen, H.).
Gyldendalske Boghandel, Nordisk forlag, København. pp. 23-30.
- Andersen, P. (1958). Fonemsystemet i Østfynsk.
Schultz, København.
- Basbøll, H. (1985). Stød in modern Danish.
Folia Linguistica, vol. 19, No. 1-2. pp. 1-50.
- Basbøll, H. (1988). The modern Danish stød and phonological weight.
Certamen Phonologicum: Papers from the 1987 Cortona Phonology Meeting.
(Red. Bertinetto, P.M. & Loporcaro, M.). Rosenberg & Sellier, Torino. pp. 119-152
- Basbøll, H. (1998). Nyt om stødet i moderne rigsdansk - om samspillet mellem lydstruktur og ordgrammatik.
Danske Studier 1998, C.A. Reitzels Forlag. pp. 33-86.
- Catford, I. (1964). Phonation types: the classification of some laryngeal components of speech production.
In Honor of Daniel Jones. (Red. Abercrombie, D., Fry, D.B., MacCarthy, P.A.D., Scott, N.C., & Trim, J.L.M.) Longmans, Green & Co., London. pp. 26-37.
- Dik, S. C. (1989). The Theory of Functional Grammar. Part I: The Structure of the Clause.
Foris Publications, Dordrecht / Providence, Rhode Island.
- Esling, J.H. (1984). Laryngographic study of phonation type and laryngeal configuration.
Journal of the International Phonetic Association, vol. 14. pp. 56-73.
- Fant, G. (1960). Acoustic theory of speech production.
Mouton, Haag.
- Fant, G. (1993). Some problems in voice source analysis.
Speech Communication, vol. 13, No. 1. pp. 7-22.
- Fant, G. (1995). The LF-model revisited. Transformations and frequency domain analysis.
STL-QPSR, vol. 36, No. 2-3. pp. 119-156.
- Fant, G., Liljencrants, J. & Lin, Q. (1985). A four-parameter model of glottal flow.
STL-QPSR, vol. 26, No. 4. pp. 1-13
- Fant, G. & Lin, Q. (1987). Glottal source - vocal tract acoustic interaction
STL-QPSR, vol. 28, No. 1. pp. 13-27.
- Fischer-Jørgensen, E. (1955). Om vokallængde i dansk rigsmål.
Nordisk Tidsskrift for Tale og Stemme 15. pp. 33-56.
- Fischer-Jørgensen, E. (1974). A note on laryngeal activity in the Danish "stød".
Status Report on Speech Research Haskins Laboratories, 39/40. pp. 255-259.

- Fischer-Jørgensen, E. (1987). A Phonetic Study of the Stød in Standard Danish
ARIPUC 21. Københavns Universitet, København. pp. 55-265.
- Fischer-Jørgensen, E. (1989a). A Phonetic Study of the Stød in Standard Danish
University of Turku, Turku.
- Fischer-Jørgensen, E. (1989b). Phonetic analysis of the stød in standard Danish.
Phonetica 46. pp. 1-59.
- Forchhammer, V. (1954). Stemmedannelse.
Nordisk Lærebog for Talepædagoger (Bind1). (Red. Blegvad, N.R., Forchhammer, V.,
Selmer, E.W., Henningsson, H.). Rosenkilde og Bagger Forlag, København. pp. 174-
199
- Frøkjær-Jensen, B. (1960). Le trans Pitchmeter
Centre universitaire de l'État à Mons - Laboratoire de phonétique, Mons.
- Faaborg-Andersen, K. (1957). Electromyographic Investigation of Intrinsic Laryngeal
Muscles in Humans.
Møllers Bogtrykkeri, København.
- Gerratt, B.R. & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation.
Journal of Phonetics, vol. 29. pp. 365-381.
- Gordon & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview.
Journal of Phonetics, vol 29. pp. 383-406.
- Grønnum, N. (1992). The Groundworks of Danish Intonation.
Museum Tusculanum Press, København.
- Grønnum, N. (1998). Danish.
Journal of the International Phonetic Association, vol. 28, No. 1&2. pp. 99-105.
- Grønnum, N. (2001). Fonetik og Fonologi (Anden udgave).
Akademisk Forlag, København.
- Grønnum, N. (2005). Fonetik og Fonologi (Tredje udgave).
Akademisk Forlag, København.
- Grønnum, N. (2007). Rødgrød med fløde : en lille bog om dansk fonetik.
Akademisk Forlag, København.
- Grønnum, N. & Basbøll, H. (2001a). Consonant length, stød and morae in Danish.
Papers from Fonetik 2001, held at Örenäs, May 30 - June 1, 2001. (Red. Karlsson, A.,
Weijer, J.v.d.,) Working Papers 49, Department of Linguistics, Lund University.
pp. 46-49.
- Grønnum, N. & Basbøll, H. (2001b). Consonant length, stød and morae in Standard Danish.
Phonetica 58, 2001, 230-253.

- Grønnum, N. & Basbøll, H. (2002a). Stød and Length: Acoustic and Cognitive Reality? Proceedings of the 1st International Conference on Speech Prosody; Aix en Provence, France, 11-13 April 2002. (Ed. Bel, B. & Martin, I.). Aix-en-Provence : Laboratoire Parole et Langage, 2002. pp. 355-358.
- Grønnum, N. & Basbøll, H. (2002b). Stød and Vowel Length: Acoustic and Cognitive Reality? Proceedings from Fonetik 2002, Fysikcentrum, Stockholm, May 29-31, 2002. (Ed. Dungen, C., Granström, B., House, D., Öster, A.M.,) TMH-QPSR Vol. 44, 2002. pp. 85-88.
- Grønnum, N. & Basbøll, H. (2003a). Two-phased stød vowels - a cognitive reality? Proceedings from Fonetik 2003, Löfvånger 2-4 June 2003. (Ed. Strangert, E., Heldner, M., Karlsson, F., Schaeffler, F., Wretling, P.) Phonum 9 (Reports in Phonetics), Umeå University, Umeå. pp. 33-36.
- Grønnum, N. & Basbøll, H. (2003b). Stød and Length in Standard Danish: Experiments in Laboratory Phonology. Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona 3-9 August 2003 (Ed. Solé, M.J., Recasens, D. & Romero, J.). Universitat Autònoma de Barcelona, Barcelona. pp. 455-458.
- Grønnum, N. & Basbøll (2007). Danish Stød - Phonological and Cognitive Issues. Experimental Approaches to Phonology. (Ed. Solé, M.J., Beddor, P.S., Ohala, M.). Oxford University Press, Oxford. pp. 192-206.
- Grønnum, N., Vazquez-Larruscaín, M. & Basbøll, H. (2013). Danish Stød: Laryngealization or Tone. *Phonetica* 2013, vol. 70. pp. 66-92.
- Hanson, H.M., Stevens, K.N., Kuo, H.K.J., Chen, M.Y., Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, vol. 29. pp. 451-480
- Johnson, K. (1997). *Acoustic & Auditory Phonetics*. Blackwell Publishers Inc., Cambridge, Massachusetts.
- Kyst, B. (2008). Trykgruppens toner i århusiansk regiolekt. *Danske talesprog*, bind 9. C.A. Reitzels forlag, København.
- Ladefoged, P. (1962). *Elements of Acoustic Phonetics*. Oliver and Boyd, Edinburgh and London.
- Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics*. The University of Chicago Press, Chicago and London.
- Laver, J. (1980). *The phonetic Description of Voice Quality*. Cambridge Studies in Linguistics 31. Cambridge University Press, Cambridge.

- Lindblom, B. (1982). The interdisciplinary challenge on speech motor control.
Speech Motor Control (Red. Grillner, S., Lindblom, B., Lubker, J. & Persson, A.).
Wenner-Gren Symposium Series, Vol 36. pp. 3-18.
- Lindqvist, J. (1969). Laryngeal mechanisms in speech.
STL-QPSR, vol. 10, No. 2-3. pp. 26-32
- Lindqvist (1972a). A descriptive model of laryngeal articulation in speech.
STL-QPSR, vol. 13, No. 2-3. pp. 1-9
- Lindqvist (1972b). Laryngeal articulation studied on Swedish subjects.
STL-QPSR, vol. 13, No. 2-3. pp. 10-27
- Maryn, Y., Roy, N., Bodt, M.D., Cauwenberge, P.V. & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis.
JASA, vol. 126, No. 5. pp. 2619-2634.
- Petersen, N.R. (2001). Akustisk Baseret Stemmebeskrivelse.
Dansk Audiologopædi, vol 37, Nr. 1. pp. 15-27.
- Petersen, N.R. (2008). Glottisfunktionens Akustik.
Upubliceret manuskript. Københavns Universitet, København.
- Petersen, P.R. (1973). An instrumental investigation of the Danish "stød"
ARIPUC 7. Københavns Universitet, København. pp. 195-234.
- Ringgaard, K. (1960). Vestjysk stød.
Universitetsforlaget i Aarhus, Århus.
- Roach, P.J. & Hardcastle, W.J. (1979). Instrumental measurement of phonation types: A laryngographic contribution.
Current Issues in the Phonetic Sciences, Current Issues in Linguistic Theory, vol. 9,
part I. (Red. Hollien, H. & Hollien, P.) John Benjamins B.V., Amsterdam.
- Schachtenhaufen, R. (2013). Fonetisk reduktion i dansk.
Ph.d.-afhandling, Københavns Universitet.
- Slifka, J. (2006). Some Physiological Correlates to Regular and Irregular Phonation at the End of an Utterance.
Journal of Voice, Vol. 20, No. 2.
- Smith, S. (1944). Bidrag til Løsning af Problemer vedrørende Stødet i Dansk Rigssprog.
Kaifers Boghandel, København.
- Stevens, K. N. (2000). Acoustic Phonetics.
The MIT Press, Cambridge, Massachusetts & London, England.
- Sundberg, J. & Gauffin, J. (1978). Waveform and spectrum of the glottal voice source.
STL-QPSR vol 19, No. 2-3. pp. 35-50

Thorsen, N.G. (1974). Acoustical and perceptual properties of the Danish stød.
ARIPUC 8, Københavns Universitet, København. pp. 207-213.

Welleschik, B. (2007). "Larynx normal".
http://commons.wikimedia.org/wiki/File:Larynx_normal.jpg

Dansk resume

Denne afhandling omhandler en akustisk-fonetisk undersøgelse af de ændringer i stemmekvaliteten der forekommer i forbindelse med stød.

Indledningsvis gives en gennemgang af de væsentligste nyere instrumentalfonetiske undersøgelser af stød. Der foreligger en betydelig base af akustiske og artikulatoriske/fysiologiske data om stød, men da realisationen af stød er præget af stor variation, er det ikke entydigt fastslået hvordan stød skal beskrives fonetisk. Det mest robuste træk ved stød er et dyk eller fald i intensitet som kan kobles til glottal indsnævring. Prototypiske stød er desuden karakteriseret ved en iøjnefaldende fase med uregelmæssige stemmelæbesvingninger der kan betegnes som (en slags) knirk. I nyere fremstillinger er dette træk blevet brugt definerende: Stød er (en slags) knirk.

Imidlertid er knirk ikke entydigt knyttet til stød: Det er velkendt at distinkte og tydelige stød kan realiseres uden uregelmæssige stemmelæbesvingninger, og uregelmæssige stemmelæbesvingninger forekommer ofte uden at det har nogen sammenhæng med stød. Forekomsten af uregelmæssige stemmelæbesvingninger er altså hverken tilstrækkelig eller nødvendig for realisationen af stød.

I afhandlingen foreslås en model for stød som kan forklare den labile tilstedeværelse af uregelmæssige stemmelæbesvingninger i stød. Hypotesen er at stød skal ses om et dynamisk stemmekvalitetsforløb der indebærer en (passende stor) ændring af stemmekvaliteten i retning af en mere komprimeret, eventuelt knirket, stemme som udfolder sig inden for et passende (kort) tidsrum.

Med henblik på at teste den opstillede hypotese udvikles en metodik der kombinerer en række akustiske analyser der kan belyse de forventede ændringer i stemmekvalitet. Hypotesen afprøves med den udviklede metode på oplæst materiale, dvs. laboratorietale, fra en enkelt informant.

Undersøgelsen viser at ændringer i stemmekvaliteten kan forløbe på en måde der er i modstrid med den opstillede hypotese, idet kompressionen ikke nødvendigvis er højest der hvor de uregelmæssige svingninger forekommer. Dermed udfordres også de bagvedliggende modeller for stemmekildens funktion idet forekomsten af uregelmæssige stemmelæbesvingninger antages at være en konsekvens af stigende kompression.

Videre analyser af målingerne afslører en polarisering i resultaterne som indikerer at stød kan realiseres på endnu to forskellige måder der begge indebærer uregelmæssige svingninger: én hvor kompressionen vurderes til at være højest sammenfaldende med eller meget nær den fase hvor de uregelmæssige stemmelæbesvingninger forekommer, og en anden måde hvor kompressionen vurderes til at være højest noget før de uregelmæssige stemmelæbesvingninger forekommer.

Det henstår til fremtidig undersøgelse at afklare om de to stødtyper kan genfindes i et større materiale og i givet fald undersøge distributionen af de to typer.

Summary in English

This thesis concerns an acoustic phonetic investigation of the changes in voice quality that may appear in relation to the Danish syllable prosody stød.

An overview of the most recent instrumental phonetic studies on the stød is presented. A substantial amount of knowledge about acoustic, articulatory and physiological properties of stød has been accumulated but as the realization is seen to be very variable, it has not yet been possible to determine how stød should properly be described phonetically. The most robust acoustic trait of the stød is a drop in intensity that can be linked to a glottal constriction (which in contrast to e.g. glottal stops in English appears to almost never be complete). Prototypical realizations of stød are characterized by a stretch of irregular vocal fold vibrations, which may be labeled as (a kind of) creak. This characteristic has in recent publications been treated as a defining feature: Stød is (a kind of) creaky voice.

However, creak is not unequivocally linked to stød: It is well known that distinct and clear realizations of stød need not exhibit irregular vocal fold vibrations, and irregular vocal fold vibrations frequently occur without having any connection to stød (which holds true for Danish, as well as for other languages). The occurrence of irregular vocal fold vibrations is thus neither sufficient nor necessary for the realization of stød.

A model for the stød is proposed which can explain the inconsistent occurrence of irregular vocal fold vibrations in stød. The hypothesis is that stød can be explained as a dynamic voice quality movement in the form of a suitably large change in phonation type in the direction of more compressed, potentially creaked, voice and back which unfolds over a suitably short time span.

A framework for testing the hypothesized model for stød is developed: It combines a range of acoustic analyses meant to capture the expected changes in voice quality. The hypothesis is tested on laboratory speech material from one subject.

The results of the experiments show that the voice quality measures may follow patterns that are not in accord with the stated hypothesis, in that the compression need not peak in conjunction with the irregular vocal fold vibrations. Based on the findings the proposed stød hypothesis is rejected. The results also challenge the underlying models of voice quality, since irregular vibrations are assumed to be the consequence of rising compression above a certain level.

Further analysis of the data reveals patterns that seem to indicate that stød may be realized in yet two different ways, that both include irregular vocal fold vibrations: One in which the compression is judged to be at its maximum close to or coinciding with the stretch of irregular vocal fold vibrations, and one in which the compression is judged to be at its maximum well before the vibrations become irregular.

It remains for future studies to confirm or disconfirm the general occurrence of the two types and if confirmed to assess if there is a distributional pattern for them.

Stød og stemmekvalitet

*En akustisk-fonetisk undersøgelse
af ændringer i stemmekvaliteten
i forbindelse med stød*

A P P E N D I X

Indhold

A Oversigt over lydfiler.....	202
B Praat script.....	203
B.I Målescript.....	203
B.II Centrale dele af målescriptet.....	215
C Afprøvning af filterbåndbredder ifht. til best. af $H_1:H_2$	218
D Test af $H_1:H_2$'s følsomhed for fejl i f_0 -bestemmelsen.....	224
E Test af BED's følsomhed for fejl i f_0 -bestemmelsen.....	230
F Analyse (kurveblade) for testfiler.....	236
G Testord.....	245
H Kurveblade for udvalgte testord.....	247
H.I ma'ler.....	248
H.II maler.....	253
H.III ma'ser.....	257
H.IV mases.....	261
H.V ta'ber.....	266
H.VI taber.....	271
H.VII li'men.....	276
H.VIII limet.....	281
H.IX lamm'et.....	285
H.X lammet.....	290

A Oversigt over lydfiler

Lydfiler

A.I.1	[i] med varierende kompression
A.I.2	[a] med varierende kompression
A.II.1	[i] med f_0 -sweep
A.II.2	[a] med f_0 -sweep
A.III.1	[i] til [a] (fortungerække)
A.III.2	[u] til [ɒ] til [a] (bagtungerække)
A.III.3	[i y u] (snævre vokaler)
A.IV.1	knirk-og-diplophoni
A.V.1	Vokal med indspilede sekvenser med højere kompression
A.VI.1	Syntetiseret vokal med amplitudemodulation
A.VI.2	Syntetiseret vokal med frekvensmodulation
A.VI.3	Syntetiseret vokal med amplitude- og frekvensmodulation
A.VII.1	mo'ser + moser : hurtig og langsom

B Praat script

B.I Målescript

```
# Last revised: 2013 05 13 using Praat v 5.3.27
# Developed using Praat v 4.6.21
# Earlier versions developed using Praat v 4.3.36
#
# 2007 07 modificering af v 1.1 for at tilpasse output til R SSANOVA-analyse
# 2007 08 yderligere modificering for tilpasning til R SSANOVA-analyse: udskrift i fire separate filer
# 2007 09 17 tilføjelse af formantestimeringer - skrives ikke til fil endnu
# 2007 09 18 gem picture som EPSfil
#       v 1.4 ny fordeling af analyser i de tre frames
#       formantværdier skrives til fil
# 2007 12 05 ændret margener (Select outer viewport... 0.3 7.5 xx xx) så det sværtede er inden for
#       hp-printerens reelle printbare område
# 2008 02 18 Modificering for at visualisere formanters båndbredde ved plotområde
# 2008 03 25 Modificering for at tillade analyse uafhængig af TextGrid og valgfri output af måledata i
#       (fil og TextGrid) eller ej
# 2008 04 03 Bandwidths skrives nu til fil.
#       Segment labels skrives nu på wb spektrogram
#       Graph_only virker ikke! Der skrives til filer og indsættes points i TextGrid ..... FIX LATER
#       v 1.5.1
#       Rettet fejl i F1 plot når b1 overstiger 300 Hz
# 2008 04 16 Tilføjet procedure som skriver kommentarer fra tier 4-9+10 ind hhv over NB spektrogram og under WB
spektrogram.
#       Valgmuligheden "include classification" er ikke implementeret..... FIX LATER
# 2013 04 23 Grafisk output justeret (Font & garniture). Vendt fortegn på dB-skalaen i ramme II i overensst. m. teksten.
# 2013 05 13 Pitch-check aktiveret: 5% f0-diff eller HNR<-10dB => H1:H2 grey, BED grey
# 2013 05 17 overvej om To Harmonicity (cc)... 0.005 50 0.1 1 skal justeres ned til 40 Hz eller 30 Hz
#       HNR falder kraftigt ifbm diplophoni hvis f0/2 falder under grænsefrekvensen på 50 Hz
#
```

form Stød analyse

```
word Filename File
sentence Target_label *
sentence Output_directory c:\praat_temp\stoedanalyse_test\
boolean Only_analyze_selected_interval 0
boolean Graph_only 0
comment Include auditory classification from tier 4-10
boolean Include_classification 1
comment
positive Labels_is_in_tier 1
positive Segments_is_in_tier 2
positive Output_in_point_tier 3
positive Margin 0.5
positive Number_of_analysis_points_per_segment 20
comment Show analyses:
boolean Show_pitch 1
boolean Intensity(0-1000) 1
boolean HNR 1
boolean H2-H1 1
boolean H1.5-H1 1
boolean CoG(0-1000) 1
boolean BED(0-1.5xpitch-1000) 1
boolean Formants 1
comment Output is appended to file in c:\praat_temp\stoedanalyse_test\'Filename'
endform
```

```

echo filename: 'filename$'
printline target label: 'target_label$'

if only_analyze_selected_interval=1

    label$ = target_label$
    labelcounter = 0

    editor LongSound 'filename$'
    start = Get start of selection
    end = Get end of selection
    endeditor

    excerpt_start = start - margin
    excerpt_end = end + margin

    # Sound version:
    # select Sound 'filename$'
    # Extract part... 'excerpt_start' 'excerpt_end' Rectangular 1.0 yes

    # LongSound version:
    select LongSound 'filename$'
    Extract part... 'excerpt_start' 'excerpt_end' yes

#   Rename... 'target_label$' 'current_label_interval'

    call analyze
    sample_point = start

    repeat
        sample_point = sample_point + 0.001
        call measure
    until sample_point > end

    Select outer viewport... 0.3 7.5 0 9.5
    Write to EPS file... 'Output_directory$' 'filename$' 'target_label$' .EPS
    call clean-up
else

    select TextGrid 'filename$'
    number_of_intervals = Get number of intervals... 'labels_is_in_tier'
    current_label_interval = 1
    labelcounter = 0

    repeat
        select TextGrid 'filename$'
        label$ = Get label of interval... 'labels_is_in_tier' 'current_label_interval'
        if label$ = target_label$
            labelcounter = labelcounter + 1
            printline match found in interval: 'current_label_interval'
            start = Get starting point... 'labels_is_in_tier' 'current_label_interval'
            end = Get end point... 'labels_is_in_tier' 'current_label_interval'

            word_start = start
            #   Adjust selection length if duration is < 0.5 sec
            duration = end - start
            if duration < 0.5
                pad = 0.5 - duration
                halfpad = pad/2
                start = start - halfpad
                end = end + halfpad
            endif
        endif
    repeat

```

```

excerpt_start = start - margin
excerpt_end = end + margin

# Sound version:
# select Sound 'filename$'
# Extract part... 'excerpt_start' 'excerpt_end' Rectangular 1.0 yes

# LongSound version:
select LongSound 'filename$'
Extract part... 'excerpt_start' 'excerpt_end' yes

Rename... 'target_label$' 'current_label_interval'

call analyze

select TextGrid 'filename$'
if include_classification = 1
    call add_classification
endif
current_segment_interval = Get interval at time... 'segments_is_in_tier' 'start'
segment_no = 0
normalized_step = 1/number_of_analysis_points_per_segment

repeat

    segment_start = Get starting point... 'segments_is_in_tier' 'current_segment_interval'
    segment_end = Get end point... 'segments_is_in_tier' 'current_segment_interval'
    segment_label$ = Get label of interval... 'segments_is_in_tier' 'current_segment_interval'
    step = (segment_end - segment_start)/number_of_analysis_points_per_segment

    sample = segment_start
    sample_no = 1
    normalized_sample_point = segment_no + (normalized_step/2)

    call draw_segment

    if segment_label$ <> ""
        repeat
            sample_point = sample + (step/2)

            call measure

            select TextGrid 'filename$'
            Insert point... 'Output_in_point_tier' 'sample_point' 'sample_point' 'pitch' 'harmonicity'
# old vers
'banddiff_a' 'banddiff_b' 'bed2' 'cog'
            Insert point... 'Output_in_point_tier' 'sample_point' 'label$' 'labelcounter'
'normalized_sample_point' 'pitch' 'harmonicity' 'banddiff_a' 'banddiff_b' 'bed2' 'cog'
            # Output format in point tier: sample_time f0 HNR H2-H1 H1½-H1 BED CoG
# old vers
            outdata$ = ""label$ 'segment_label$' 'sample_no' 'sample_point' 'pitch:2' 'harmonicity:4'
'banddiff_a:4' 'banddiff_b:4' 'bed2:4' 'cog:4'newline$"
            if pitch_b > 2
                outdata$ =
""label$"tab$"labelcounter"tab$"normalized_sample_point:3"tab$"pitch:2"tab$0"tab$0"newline$"
                outdata$ >> 'Output_directory$'filename$'.pitch.txt
            endif

            outdata$ =
""label$"tab$"labelcounter"tab$"normalized_sample_point:3"tab$"pitch:2"tab$"cog:4"tab$0"newline$"
            outdata$ >> 'Output_directory$'filename$'.interpolated_pitch_and_CoG.txt

            outdata$ =
""label$"tab$"labelcounter"tab$"normalized_sample_point:3"tab$"banddiff_a:4"tab$"banddiff_b:4"tab$0"newline$"

```

```

        outdata$ >> 'Output_directory'$filename$.H2-H1_and_H1.5-H1.txt

        outdata$ =
        ""label$"tab$"labelcounter"tab$"normalized_sample_point:3"tab$"bed2:4"tab$"harmonicity:4"tab$'0'newline$"
        outdata$ >> 'Output_directory'$filename$.BED_and_HNR.txt

        outdata$ =
        ""label$"tab$"labelcounter"tab$"normalized_sample_point:3"tab$"f1:0"tab$"f2:0"tab$"f3:0"newline$"
        outdata$ >> 'Output_directory'$filename$.formants.txt

        outdata$ =
        ""label$"tab$"labelcounter"tab$"normalized_sample_point:3"tab$"b1:0"tab$"b2:0"tab$"b3:0"newline$"
        outdata$ >> 'Output_directory'$filename$.bandwidths.txt

# Output format in file: word segment sample_no sample_time f0 HNR H2-H1 H1½-H1 BED CoG
        sample = sample + step
        sample_no = sample_no + 1
        normalized_sample_point = normalized_sample_point + normalized_step
until sample_no > number_of_analysis_points_per_segment
endif
        current_segment_interval = current_segment_interval + 1
        segment_no = segment_no + 1
until segment_end >= end

        Select outer viewport... 0.3 7.5 0 9.5
        label$ = replace$(label$,".", "_",0)
        label$ = replace$(label$,"", "_",0)
        Write to EPS file... 'Output_directory'$filename$'.label$'.labelcounter'.EPS
        call clean-up
endif
        current_label_interval = current_label_interval+1
until current_label_interval = number_of_intervals
endif

```

```
#####

procedure analyze

#####

sound_name$ = selected$ ("Sound", 1)

Erase all
Line width... 2
Black
12
Helvetica

Select outer viewport... 0.3 7.5 0 1.5
select Sound 'sound_name$'
Draw... 'start' 'end' 0 0 no curve
Draw inner box
One mark bottom... 'start' 0 0 0 'start:2' Sec
One mark bottom... 'end' 0 0 0 'end:2' Sec

texttop$= "File: 'filename$'  Label: 'label$'  Token: 'labelcounter'"
Text top... no 'texttop$'

# nb spectrogram & 2 x pitch *****

Select outer viewport... 0.3 7.5 1 5.5

select Sound 'sound_name$'
To Spectrogram... 0.05 5000 0.002 20 Gaussian
Paint... 'start' 'end' 0 1000 100 yes 40 6 0.3 no
Draw inner box
One mark left... 0 0 0 0 Hz
One mark left... 1000 0 0 1000 Hz
One mark right... 0 0 0 30 dB
One mark right... 250 0 0 1 20 dB
One mark right... 500 0 0 1 10 dB
One mark right... 750 0 0 1 0 dB
One mark right... 1000 0 0 1 -10 dB
#Text left... 1 Frequency (Hz)
Remove

select Sound 'sound_name$'
To Pitch (ac)... 0.001 75 15 yes 0.03 0.45 0.01 0.4 0.14 600

select Pitch 'sound_name$'
Plain line
Maroon
Line width... 8
if show_pitch=1
    Draw... 'start' 'end' 0 250 no
endif

select Sound 'sound_name$'
Filter (pass Hann band)... 50 1000 100
To Pitch (ac)... 0.001 75 15 yes 0.03 0.45 0.01 0.4 0.14 300

select Pitch 'sound_name$_band'
Yellow
Line width... 3
if show_pitch=1
    Draw... 'start' 'end' 0 250 no
endif
```

```

select Pitch 'sound_name$'_band
Smooth... 20
Interpolate
#Smooth... 20
Plain line
Black
Line width... 2
if show_pitch=1
    Draw... 'start' 'end' 0 250 no
endif

Axes... start end 0 40
#Dotted line
Line width... 2
#Draw line... start 10 end 10
#Draw line... start 20 end 20
Plain line
Draw line... start 30 end 30

# wb spectrogram, harmonicity ( + intensity) *****

Select outer viewport... 0.3 7.5 5 9.5
Line width... 2

select Sound 'sound_name$'
To Spectrogram... 0.005 5000 0.002 20 Gaussian
Paint... 'start' 'end' 0 4000 100 yes 35 6 0.3 no
Draw inner box
One mark left... 0 0 0 0 0 Hz
One mark left... 4000 0 0 0 4000 Hz
One mark right... 0 0 0 1 -10 dB
One mark right... 1000 0 0 1 0 dB
One mark right... 2000 0 0 1 10 dB
One mark right... 3000 0 0 1 20 dB
One mark right... 4000 0 0 1 30 dB
#Text left... 1 Frequency (Hz)
Remove

select Sound 'sound_name$'_band
To Intensity... 75 0.005 yes

select Intensity 'sound_name$'_band
Blue
Line width... 2
if intensity=1
    Draw... 'start' 'end' 40 80 no
endif

select Sound 'sound_name$'_band
To Harmonicity (cc)... 0.005 50 0.1 1

Black
select Harmonicity 'sound_name$'_band
Line width... 3
Dotted line
if 'HNR'=1
    Draw... 'start' 'end' -10 30
endif

```

```

Line width... 2
#Draw line... start 20 end 20
#Draw line... start 10 end 10
Plain line
Draw line... start 0 end 0

# spectrogram used for spectral slices + formant analysis *****

select Sound 'sound_name$'
To Spectrogram... 0.06 500 0.002 20 Gaussian

select Sound 'sound_name$'
To Formant (burg)... 0.002 5 5000 0.025 50

#####

endproc

#####

```



```
#####

procedure measure

#####

banddiff_a = undefined
banddiff_b = undefined

Plain line
Black
Line width... 1

time = sample_point
Select outer viewport... 0.3 7.5 1 5.5

    Axes... start end -30 10
    select Pitch 'sound_name$_band
    pitch = Get value at time... 'time' Hertz linear
    if pitch = undefined
        pitch = 1
    endif
    select Pitch 'sound_name$'
    pitch_b = Get value at time... 'time' Hertz linear
    if pitch_b = undefined
        pitch_b = 2
    endif
    select Harmonicity 'sound_name$_band
    harmonicity = Get value at time... 'time' Cubic
    if harmonicity = undefined
        harmonicity = -200
    endif
    pitch_check=pitch/pitch_b

    pitch_ok = 0
    if ((pitch_check < 1.05) and (pitch_check > 0.95) and (harmonicity > -10))
        pitch_ok = 1
    endif

#   if harmonicity > -10
    if pitch > 1
        select Spectrogram 'sound_name$'
        To Spectrum (slice)... 'time'
        select Spectrum 'sound_name$'
        h1l = pitch - 10
        h1h = pitch + 10
        h2l = (pitch*2)-10
        h2h = (pitch*2)+10
        banddiff_a = Get band energy difference... h1l h1h h2l h2h

        if 'H2-H1'=1
            if pitch_ok = 0
                if (pitch = 1 or pitch_b = 2)
                    Paint circle (mm)... grey 'time' 'banddiff_a' 1
                else
                    Paint circle (mm)... black 'time' 'banddiff_a' 1
                endif
            endif
        endif

        a1l= h1l
        a1h= h1h
        a2l = (pitch*1.5)-10
        a2h = (pitch*1.5)+10
    endif
endif

###
```

```

banddiff_b = Get band energy difference... a1l a1h a2l a2h
if 'H1.5-H1'=1
    if (banddiff_b > -30 and banddiff_b < 10)
#         ## paint if there is an interpolated pitch value ##
        if pitch > 1
            Paint circle (mm)... lime 'time' 'banddiff_b' 1
        endif
    endif
endif

#     b1l= h1l*0.95
#     b1h= h1h*0.95
#     b2l= h2l*0.95
#     b2h= h2h*0.95
#     banddiff = Get band energy difference... b1l b1h b2l b2h
#     Paint circle (mm)... cyan 'time' 'banddiff' 1

Remove
endif

Axes... start end 0 200
select Spectrogram 'sound_name$'
To Spectrum (slice)... 'time'
cog = Get centre of gravity... 2.0
bed_border = pitch*1.5
bed2 = Get band energy difference... 0 'bed_border' 'bed_border' 500
if bed2 = undefined
    bed2 = 0
endif
Remove

if 'CoG'=1
    Axes... start end 0 1000
    Paint circle (mm)... Blue 'time' 'cog' 1
endif

if 'BED'=1
    Axes... start end -30 10
    if pitch > 1
#         .         .         if pitch_ok = 1
#             .         .         Paint circle (mm)... Green 'time' 'bed2' 1
#             .         .         else
#             .         .         .         Paint circle (mm)... Grey 'time' 'bed2' 1
#             .         .         .         endif
    endif
endif

Axes... start end -60 20
Plain line
Black
Draw line... start 0 end 0

Select outer viewport... 0.3 7.5 5 9.5
Axes... start end 0 4000

select Formant 'sound_name$'
f1 = Get value at time... 1 'time' Hertz Linear
b1 = Get bandwidth at time... 1 'time' Hertz Linear
f2 = Get value at time... 2 'time' Hertz Linear
b2 = Get bandwidth at time... 2 'time' Hertz Linear
f3 = Get value at time... 3 'time' Hertz Linear
b3 = Get bandwidth at time... 3 'time' Hertz Linear

```

```

if (formants=1) and (harmonicity >= -10)
    f1_low = f1-(b1/2)
    f1_hi = f1+(b1/2)
    f1_plot = f1_low
    f2_low = f2-(b2/2)
    f2_hi = f2+(b2/2)
    f2_plot = f2_low
    f3_low = f3-(b3/2)
    f3_hi = f3+(b3/2)
    f3_plot = f3_low

    if b1 < 300
        while f1_plot <= f1_hi
            Paint circle (mm)... Red 'time' 'f1_plot' 0.5
            f1_plot = f1_plot + 10
        endwhile
    elseif b1 < 500
        while f1_plot <= f1_hi
            Paint circle (mm)... Maroon 'time' 'f1_plot' 0.5
            f1_plot = f1_plot + 10
        endwhile
    elseif b1 < 1000
        while f1_plot <= f1_hi
            Paint circle (mm)... Grey 'time' 'f1_plot' 0.5
            f1_plot = f1_plot + 10
        endwhile
    endif
    if b1 < 50
        Paint circle (mm)... Yellow 'time' 'f1' 0.5
    endif

    if b1 < 1000
        if b2 < 300
            while f2_plot <= f2_hi
                Paint circle (mm)... Red 'time' 'f2_plot' 0.5
                f2_plot = f2_plot + 10
            endwhile
        elseif b2 < 500
            while f2_plot <= f2_hi
                Paint circle (mm)... Maroon 'time' 'f2_plot' 0.5
                f2_plot = f2_plot + 10
            endwhile
        elseif b2 < 1000
            while f2_plot <= f2_hi
                Paint circle (mm)... Grey 'time' 'f2_plot' 0.5
                f2_plot = f2_plot + 10
            endwhile
        endif
        if b2 < 50
            Paint circle (mm)... Yellow 'time' 'f2' 0.5
        endif
        if b2 < 1000
            if b3 < 300
                while f3_plot <= f3_hi
                    Paint circle (mm)... Red 'time' 'f3_plot' 0.5
                    f3_plot = f3_plot + 10
                endwhile
            elseif b3 < 500
                while f3_plot <= f3_hi
                    Paint circle (mm)... Maroon 'time' 'f3_plot' 0.5
                    f3_plot = f3_plot + 10
                endwhile
            elseif b3 < 1000

```

```

        while f3_plot <= f3_hi
            Paint circle (mm)... Grey 'time' 'f3_plot' 0.5
            f3_plot = f3_plot + 10
        endwhile
    endif
    if b3 < 50
        Paint circle (mm)... Yellow 'time' 'f3' 0.5
    endif
endif
endif
endif

if (b1 > 500 or b1 = undefined)
    f1=undefined
    f2=undefined
    f3=undefined
endif
if (b2 > 500 or b2 = undefined)
    f2=undefined
    f3=undefined
endif
if (b3 > 500 or b3 = undefined)
    f3=undefined
endif
if harmonicity < -10
    f1=undefined
    f2=undefined
    f3=undefined
endif

#####

endproc

#####

#####

procedure draw_segment

#####

Plain line
Black
Line width... 1

Select outer viewport... 0.3 7.5 1 5.5
Axes... start end 0 1
Draw line... segment_start 0 segment_start 1

Select outer viewport... 0.3 7.5 5 9.5
Axes... start end 0 1
Draw line... segment_start 0 segment_start 1

segment_mid = ((segment_end - segment_start)/2) + segment_start
Text... segment_mid Centre 1.05 Half 'segment_label$'

#####

endproc

#####

```

```
#####

procedure add_classification

#####

concat_comments$ = ""
for i from 4 to 9
    label_interval = Get interval at time... 'i' 'word_start'
    tier_name$ = Get tier name... 'i'
    c_label$ = Get label of interval... 'i' 'label_interval'
    concat_comments$ = concat_comments$ + tier_name$ + ": " + c_label$ + " "
endfor
Select outer viewport... 0.3 7.5 1 5.5
Text top... no 'concat_comments$'

label_interval = Get interval at time... 10 'word_start'
c_label$ = Get label of interval... 10 'label_interval'

Select outer viewport... 0.3 7.5 5 9.5
Text bottom... no 'c_label$'

#####

endproc

#####

#####

procedure clean-up

#####

select Sound 'sound_name$'
Remove
select Pitch 'sound_name$'
Remove
select Sound 'sound_name$'_band
Remove
select Pitch 'sound_name$'_band
Remove
select Intensity 'sound_name$'_band
Remove
select Pitch 'sound_name$'_band
Remove
select Pitch 'sound_name$'_band
Remove
select Formant 'sound_name$'
Remove
select Harmonicity 'sound_name$'_band
Remove
select Spectrogram 'sound_name$'
Remove

#####

endproc

#####
```

B.II Centrale dele af målescriptet

Smalbåndsspektrogram (ramme II)

select Sound 'sound_name\$'

To Spectrogram...

Window length (s) 0.05
Maximum frequency (Hz) 5000
Time step (s) 0.002
Frequency step (Hz) 20
Window shape Gaussian

Paint...

Time range (s) 'start' 'end'
Frequency range (Hz) 0 1000
Maximum (dB/Hz) 100
Autoscaling yes
Dynamic Range (dB) 40
Pre-emphasis (dB/oct) 6
Dynamic compression 0.3
Garnish no

select Sound 'sound_name\$'

Grundtone : 2 estimator + smoothing og interpolation

To Pitch (ac)...

Time step (s) 0.001
Pitch floor (Hz) 75
Max. number of candidates 15
Very accurate yes
Silence threshold 0.03
Voicing threshold 0.45
Octave cost 0.01
Octave-jump cost 0.4
Voiced/unvoiced cost 0.14
Pitch ceiling (Hz) 600

select Sound 'sound_name\$'

Filter (pass Hann band)...

From frequency (Hz) 50
To frequency (Hz) 1000
Smoothing (Hz) 100

To Pitch (ac)...

Time step (s) 0.001
Pitch floor (Hz) 75
Max. number of candidates 15
Very accurate yes
Silence threshold 0.03
Voicing threshold 0.45
Octave cost 0.01
Octave-jump cost 0.4
Voiced/unvoiced cost 0.14

```

    Pitch ceiling (Hz) 300

select Pitch 'sound_name$_band

Smooth...
    Bandwidth (Hz) 20

Interpolate

# Bredbåndsspektrogram (ramme III)

select Sound 'sound_name$'

To Spectrogram... 0.005 5000 0.002 20 Gaussian
    Window length (s) 0.005
    Maximum frequency (Hz) 5000
    Time step (s) 0.002
    Frequency step (Hz) 20
    Window shapeGaussian

Paint...
    Time range (s) 'start' 'end'
    Frequency range (Hz) 0 4000
    Maximum (dB/Hz) 100
    Autoscaling yes
    Dynamic Range (dB) 35
    Pre-emphasis (dB/oct) 6
    Dynamic compression 0.3
    Garnish no

select Sound 'sound_name$_band

# Intensitet

To Intensity...
    Minimum pitch (Hz) 75
    Time step (s) 0.005
    Subtract mean yes

select Sound 'sound_name$_band

# Harmonicity to Noise Ratio

To Harmonicity (cc)...
    Time step (s) 0.005
    Minimum pitch (Hz) 50
    Silence threshold 0.1
    Periods per window 1

```

Formantanalyse

```
select Sound 'sound_name$'
```

```
To Formant (burg)...
```

```
Time step (s) 0.002  
Max. number of formants 5  
Maximum formant (Hz) 5000  
Window length 0.025  
Pre-emphasis 50
```

Spectrogram der bruges til beregning af de spektrale mål (H1:H2, BED, CoG) + beregning af disse

```
select Sound 'sound_name$'
```

```
To Spectrogram...
```

```
Window length (s) 0.06  
Maximum frequency (Hz) 500  
Time step (s) 0.002  
Frequency step (Hz) 20  
Window shape Gaussian
```

```
select Spectrogram 'sound_name$'
```

```
To Spectrum (slice)...
```

```
Time (seconds) 'time'
```

```
select Spectrum 'sound_name$'
```

```
h1l = pitch - 10
```

```
h1h = pitch + 10
```

```
h2l = (pitch*2)-10
```

```
h2h = (pitch*2)+10
```

```
banddiff_a = Get band energy difference...
```

```
Low band floor (Hz) 'h1l'  
Low band ceiling (Hz) 'h1h'  
High band floor 'h2l'  
High band ceiling 'h2h'
```

```
select Spectrum 'sound_name$'
```

```
cog = Get centre of gravity...
```

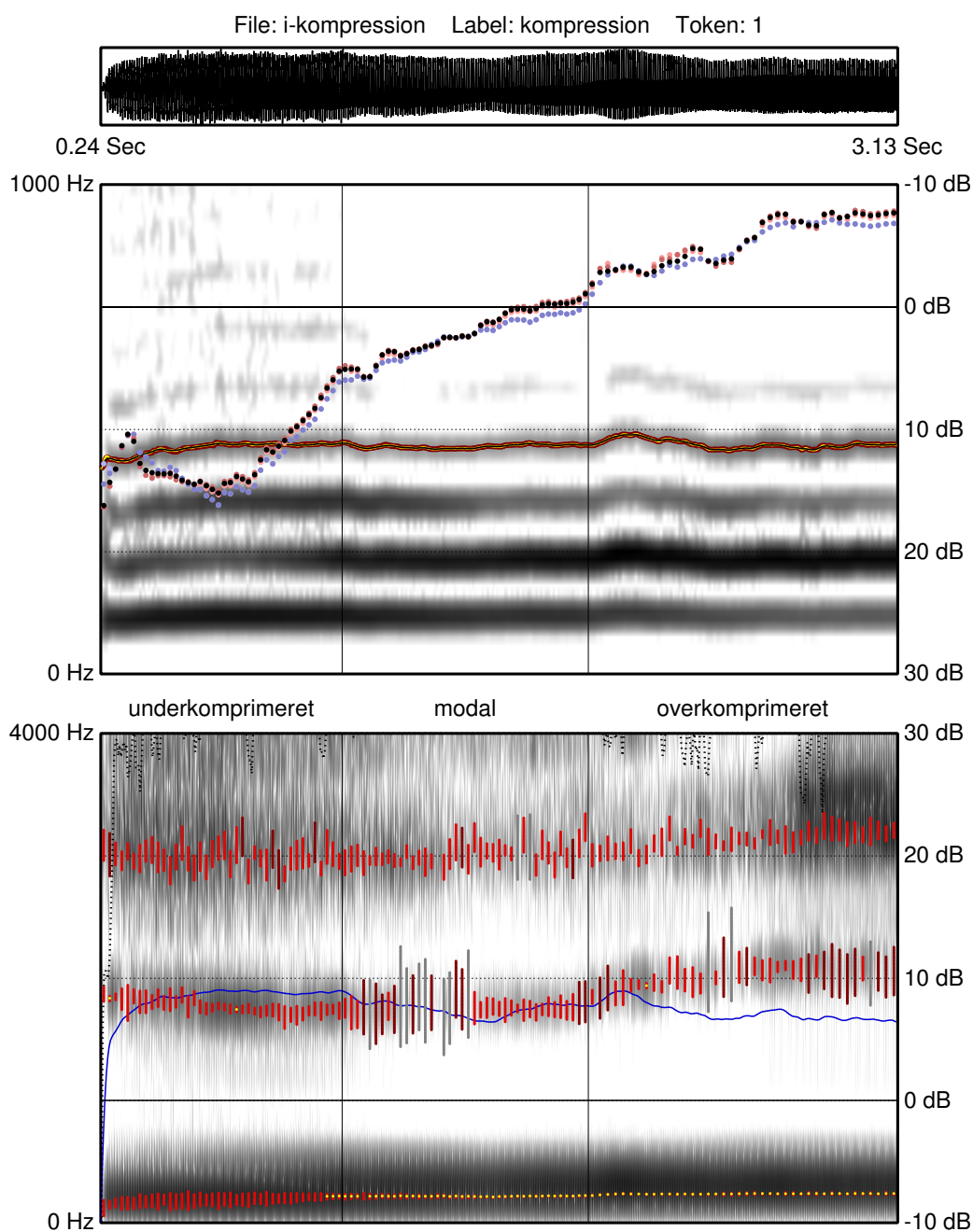
```
Power 2.0
```

```
bed_border = pitch*1.5
```

```
bed2 = Get band energy difference... 0 'bed_border' 'bed_border' 500
```

```
Low band floor (Hz) 0  
Low band ceiling (Hz) 'bed_border'  
High band floor 'bed_border'  
High band ceiling 500
```

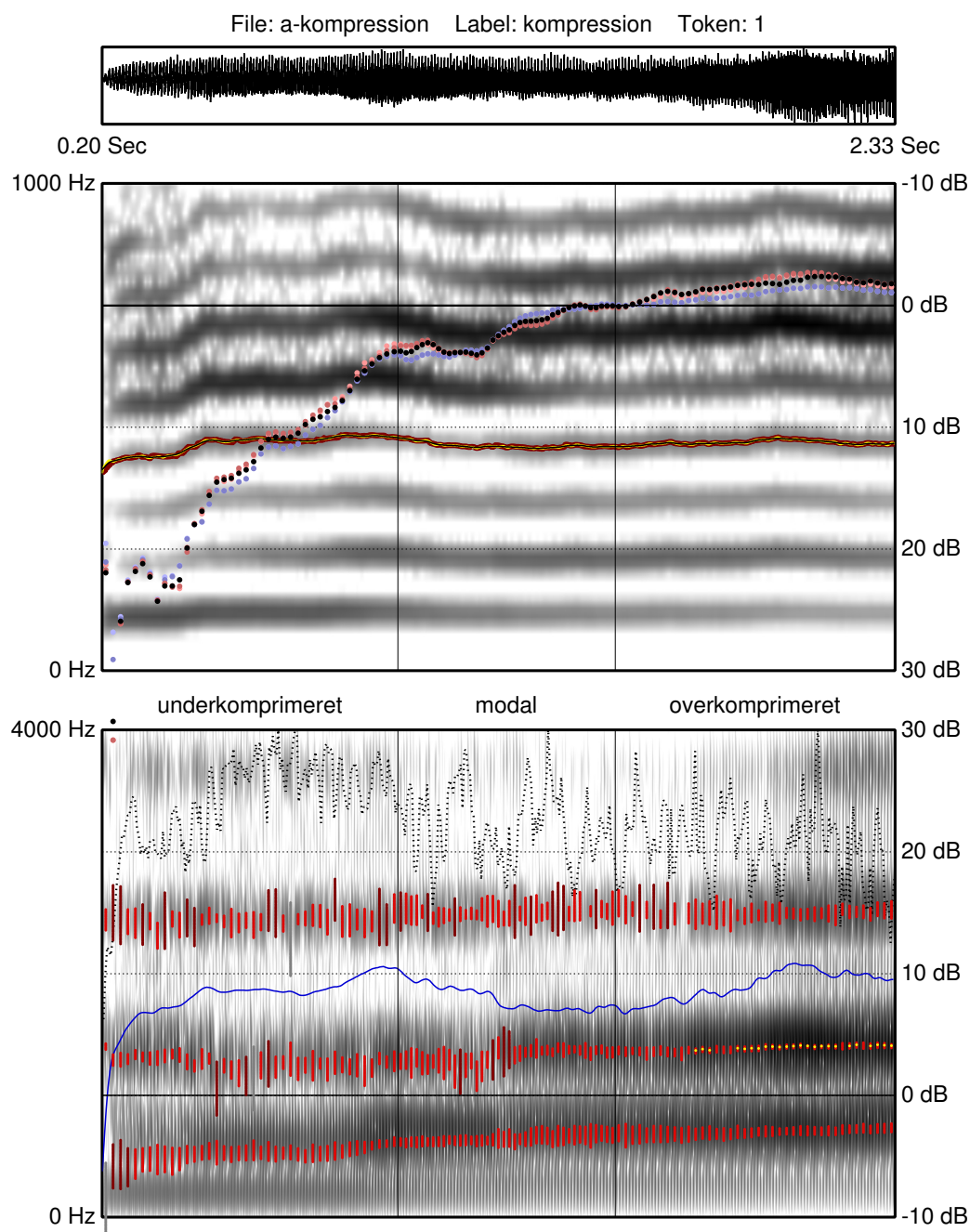

C Afprøvning af filterbåndbredder ifht. til best. af H1:H2



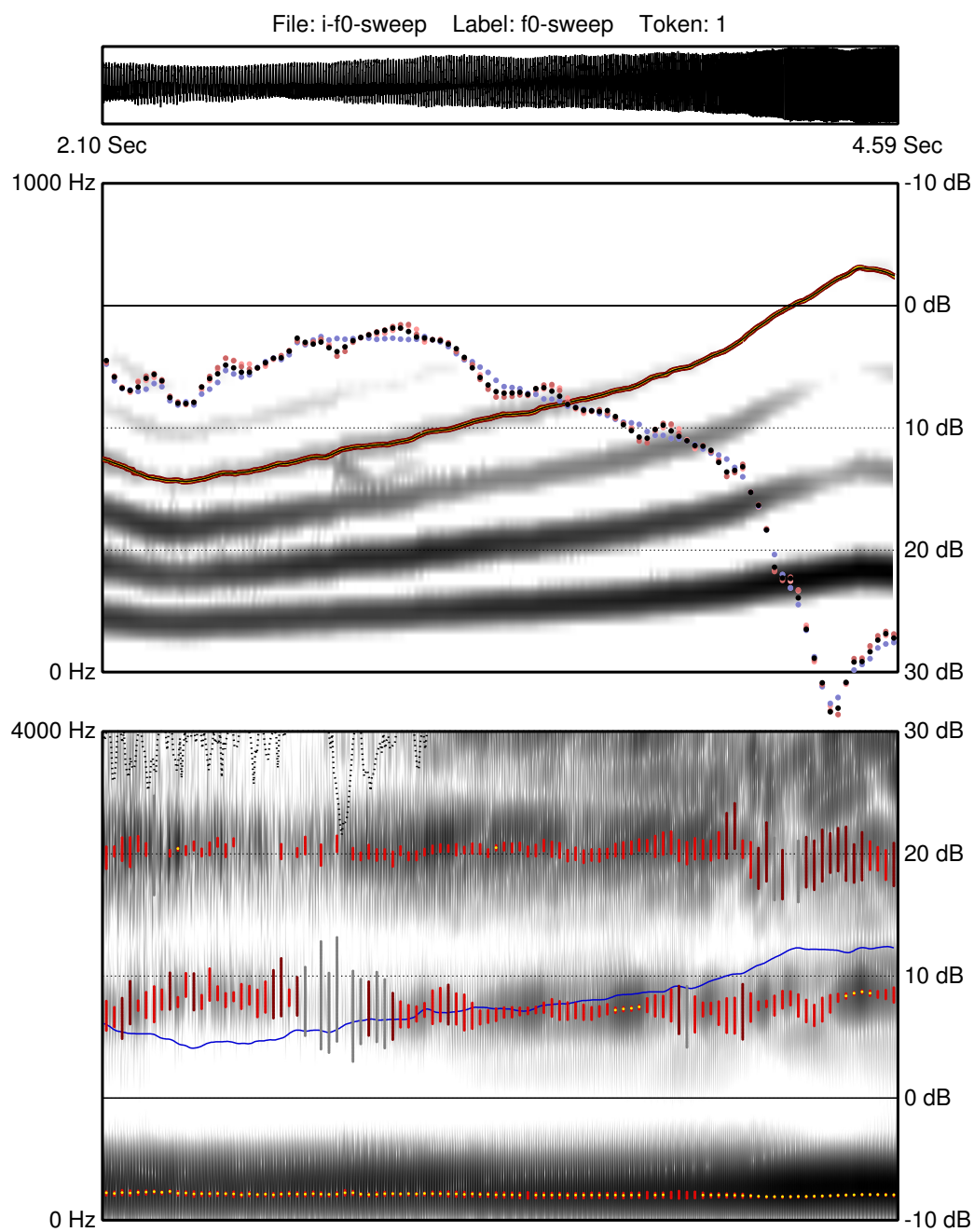
C.I. [i], varierende kompression fra underkomprimeret til overkomprimeret.

H1:H2 beregnet ud fra frekvensbånd af varierende bredde;

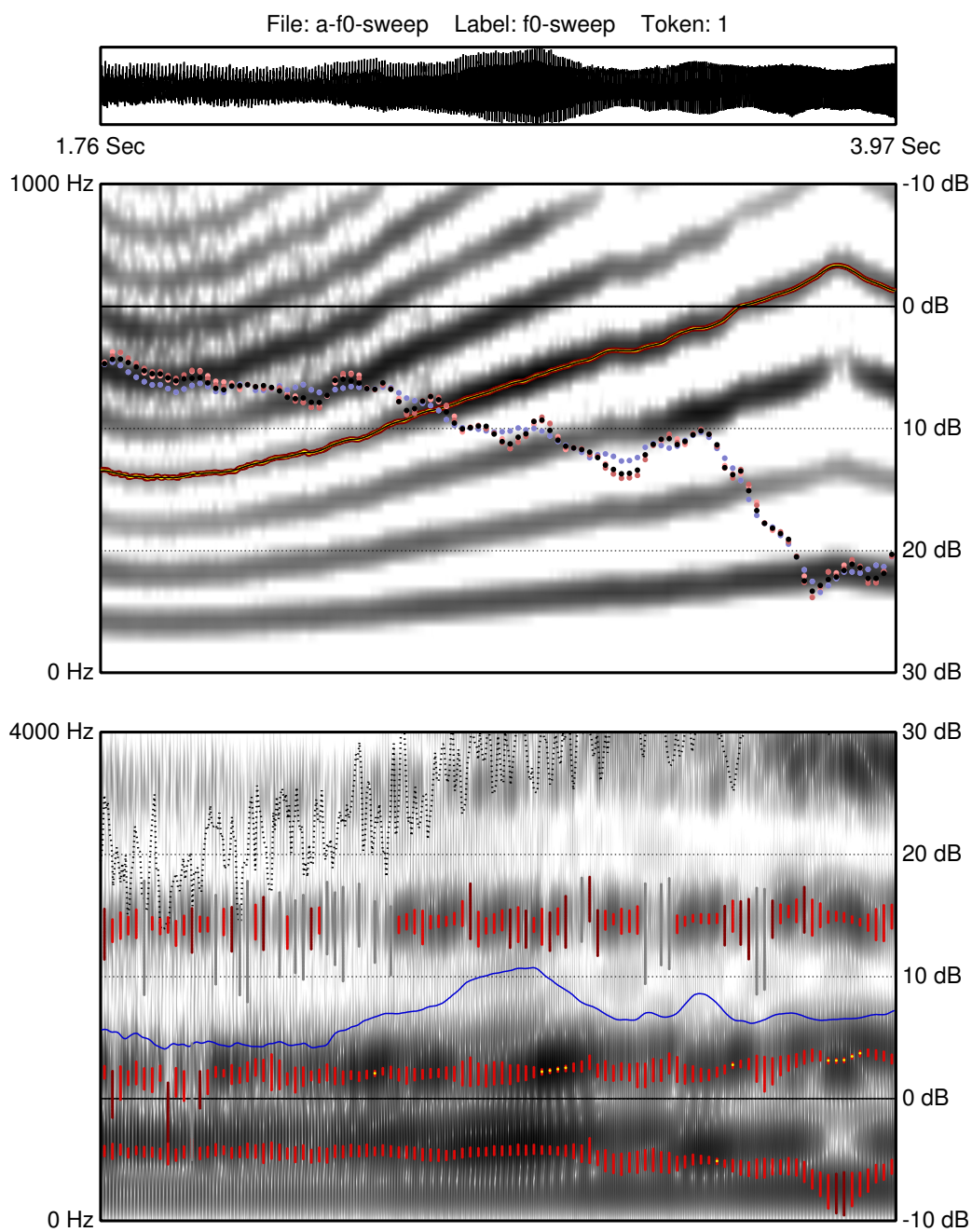
lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.



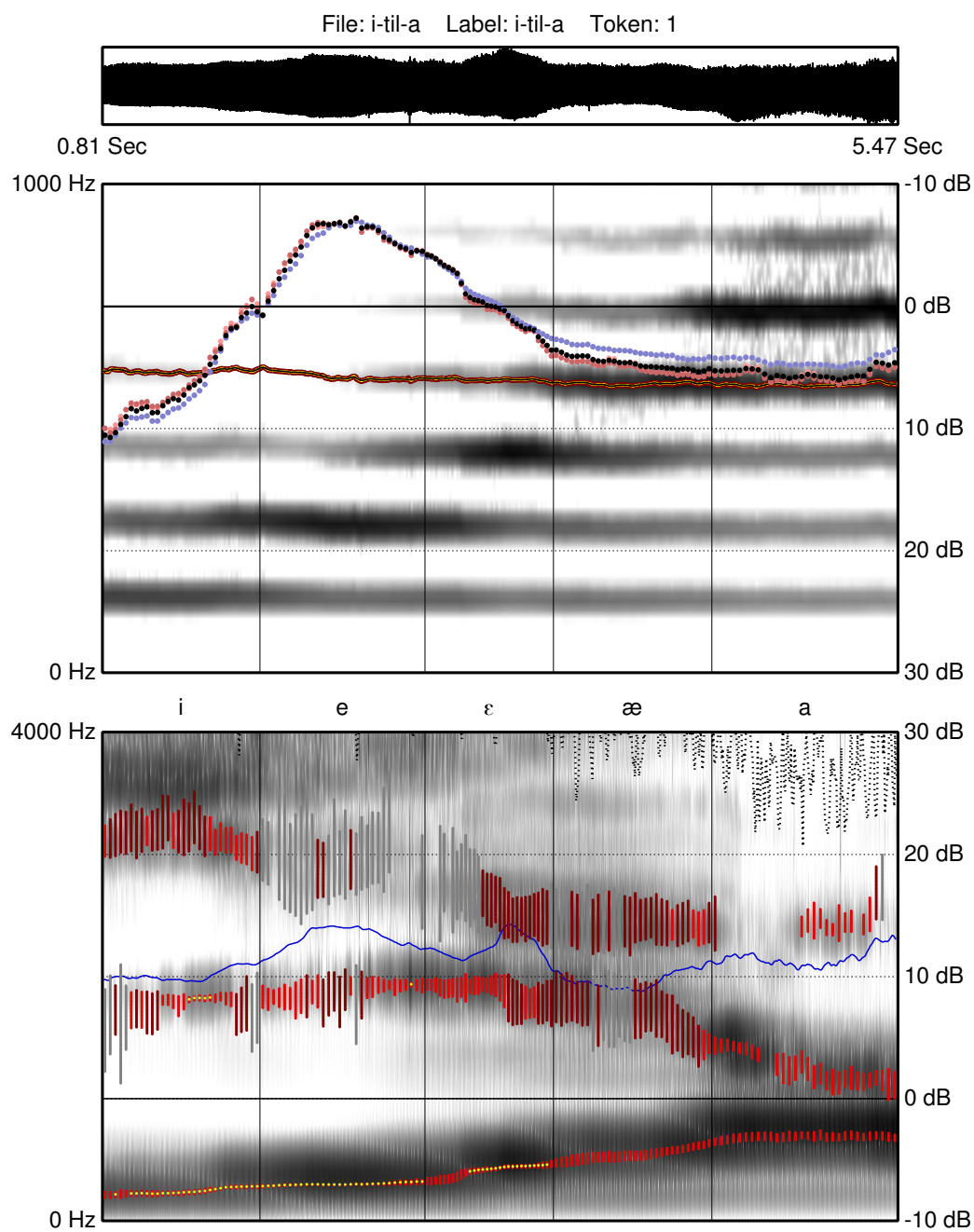
C.II. [a], varierende kompression fra underkomprimeret til overkomprimeret.
H1:H2 beregnet ud fra frekvensbånd af varierende bredde;
lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.



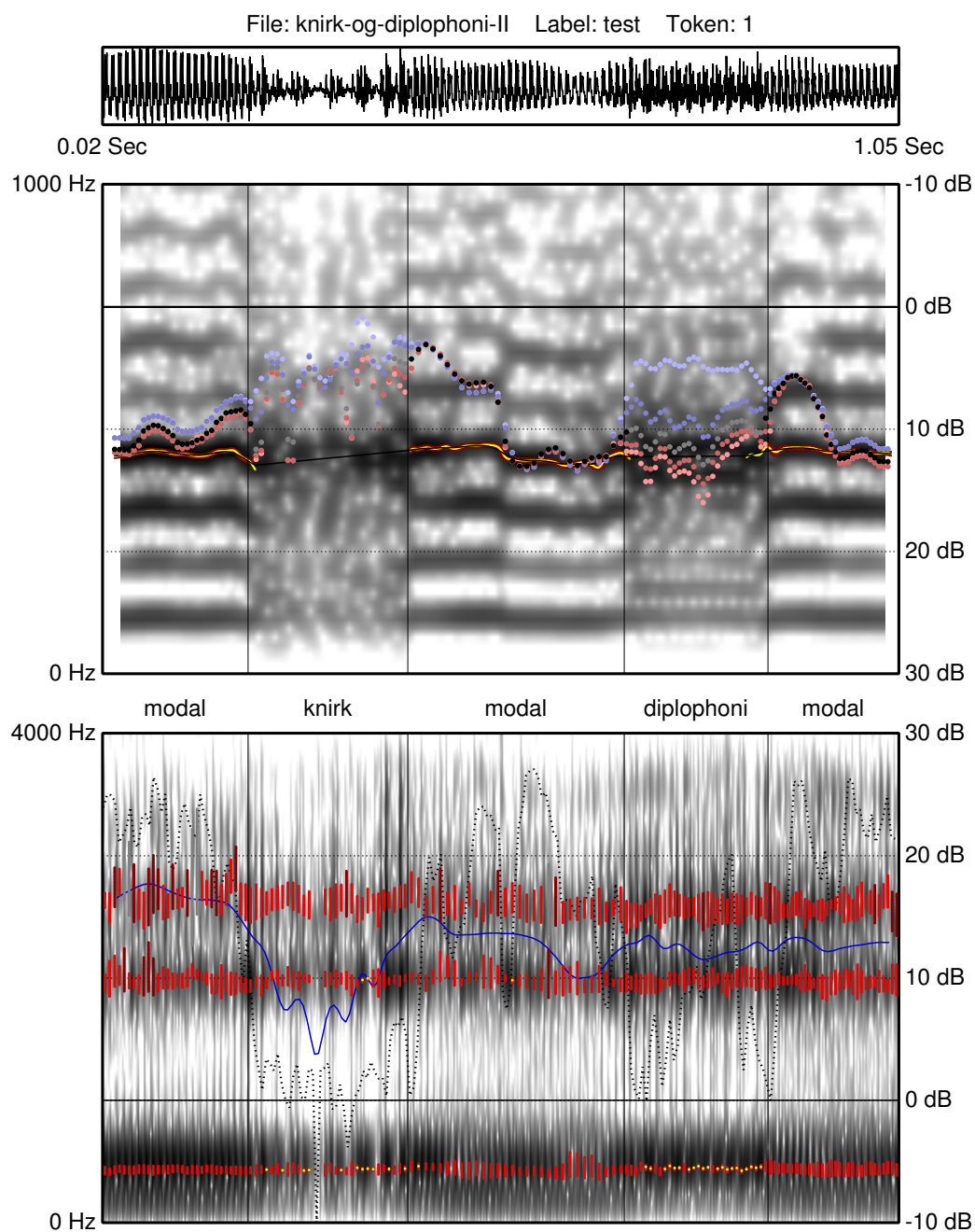
C.III. [i], f_0 -sweep. H1:H2 beregnet ud fra frekvensbånd af varierende bredde;
lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.



C.IV. [a], f_0 -sweep. H1:H2 beregnet ud fra frekvensbånd af varierende bredde;
lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.

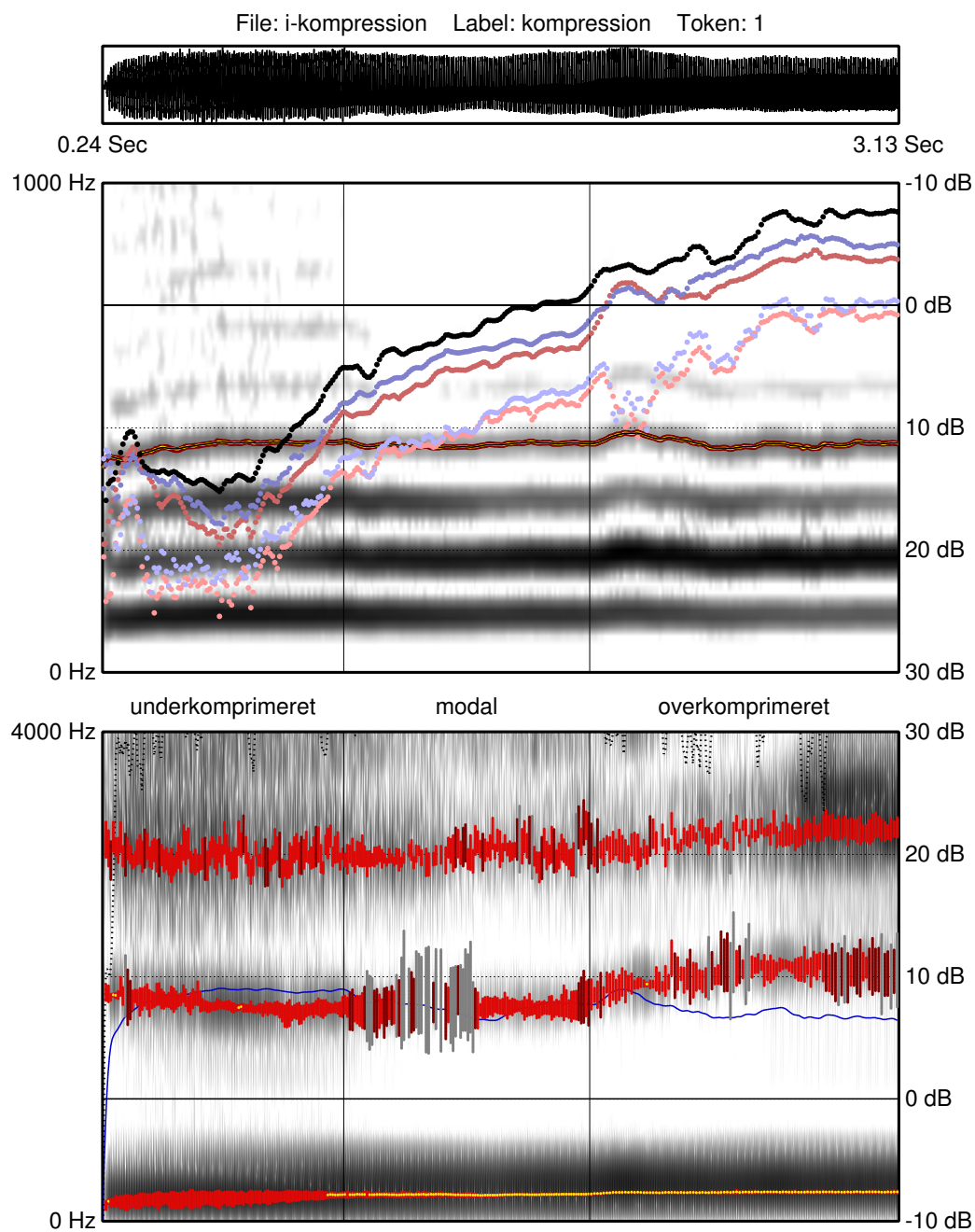


C.V. Fortunerække. H1:H2 beregnet ud fra frekvensbånd af varierende bredde;
 lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.



C.VI. [æ] med tre typer stemthed: modal, knirk og diplofoni.
H1:H2 beregnet ud fra frekvensbånd af varierende bredde;
lys blålig: 100 Hz, blålig: 50 Hz, sort: 20 Hz, rødlig: 10 Hz, lys rødlig: 5 Hz.

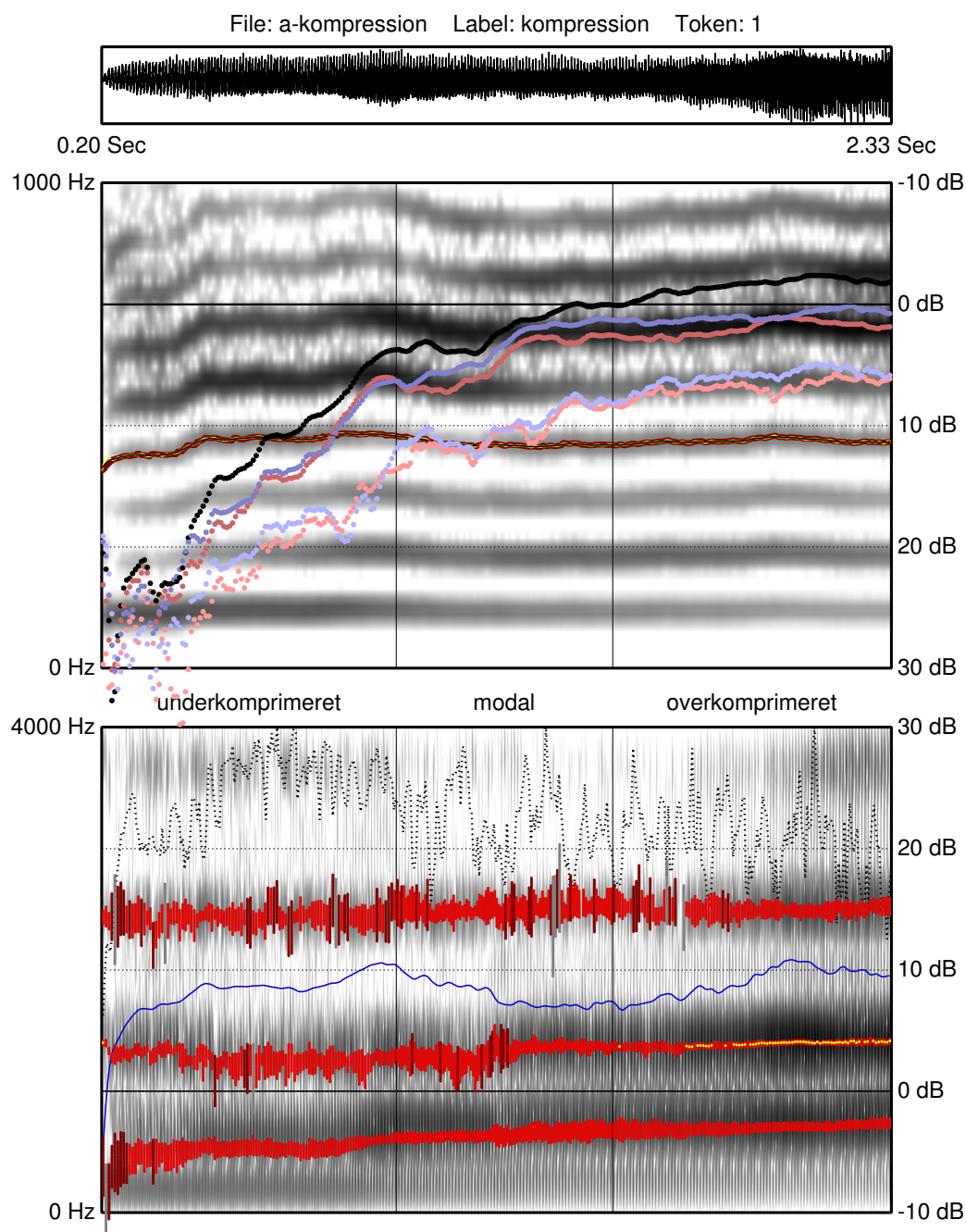
D Test af H1:H2's følsomhed for fejl i f_0 -bestemmelsen



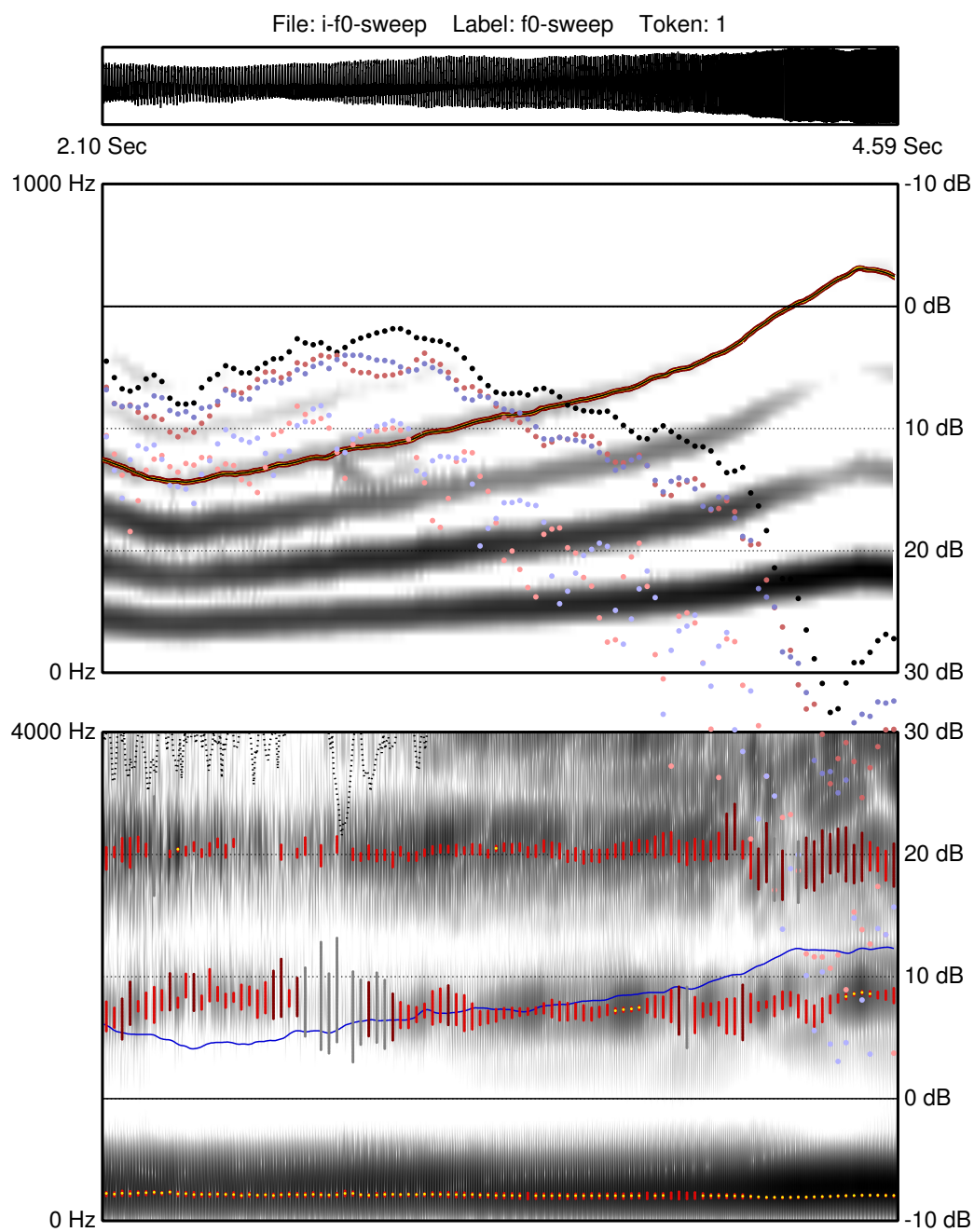
D.I. [i] med kompression varierende fra underkomprimeret til overkomprimeret.

H1:H2 beregnet ud fra f_0 med offset af forskellig grad:

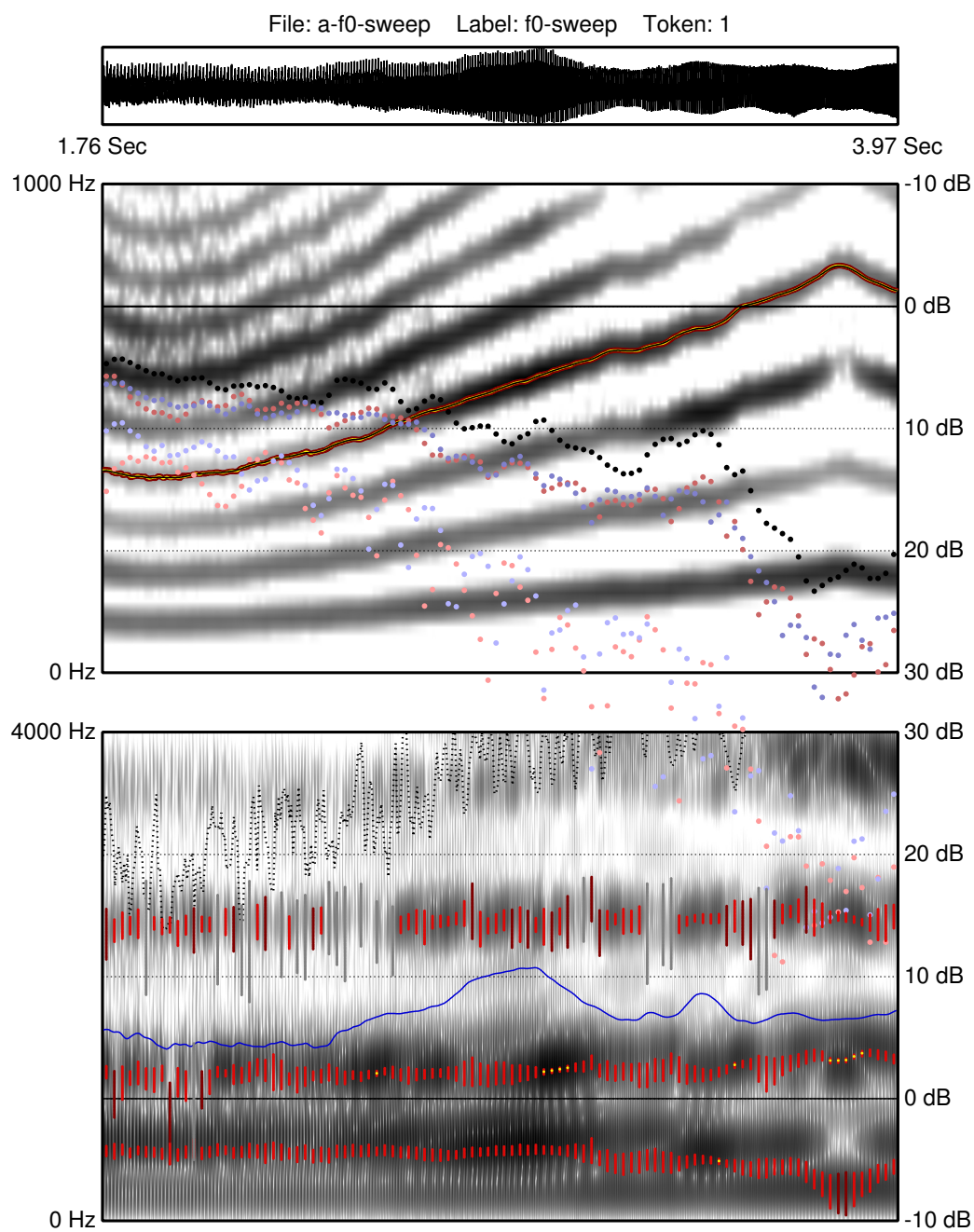
lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.



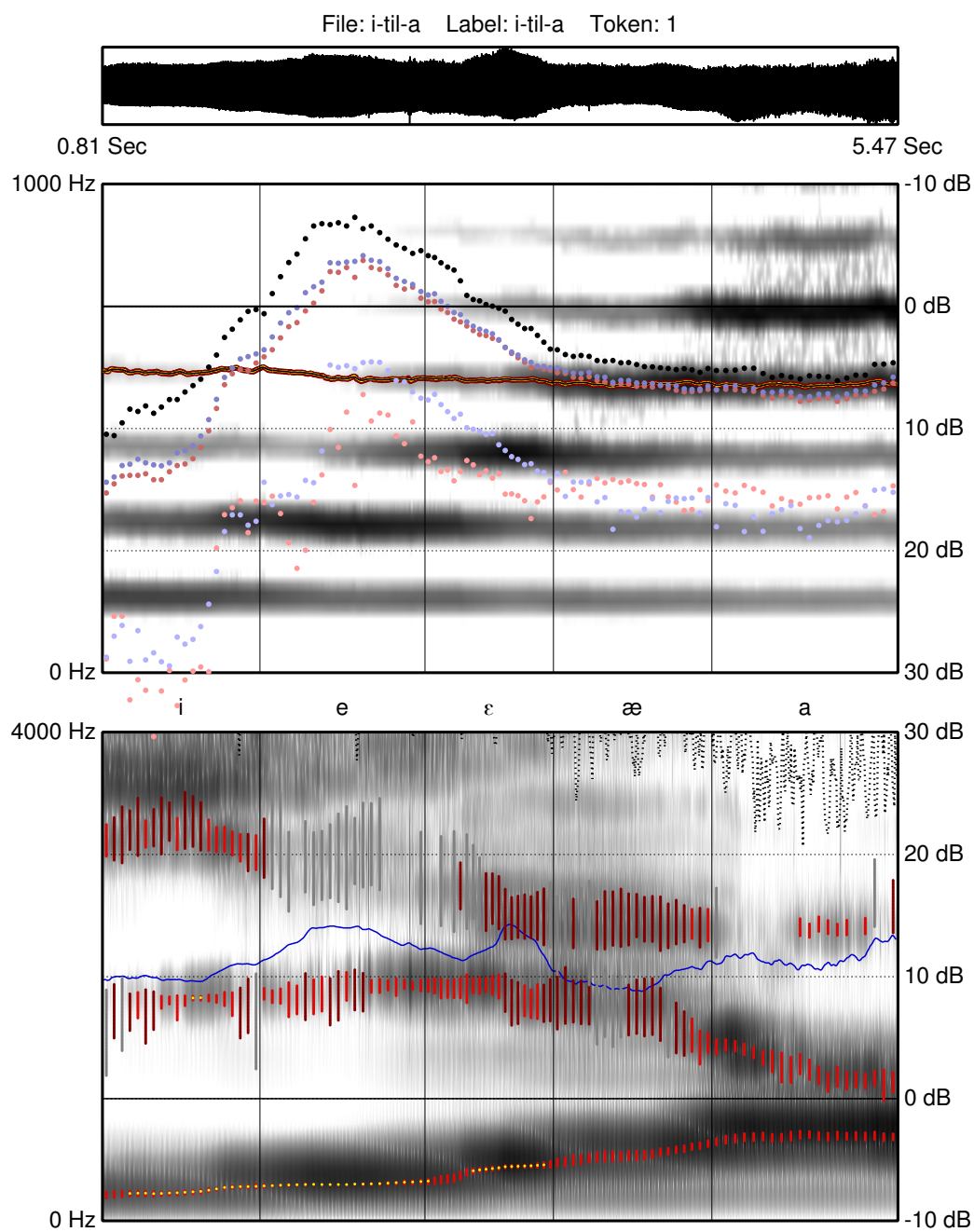
D.II. [a] med kompression varierende fra underkomprimeret til overkomprimeret.
H1:H2 beregnet ud fra f_0 med offset af forskellig grad:
lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.



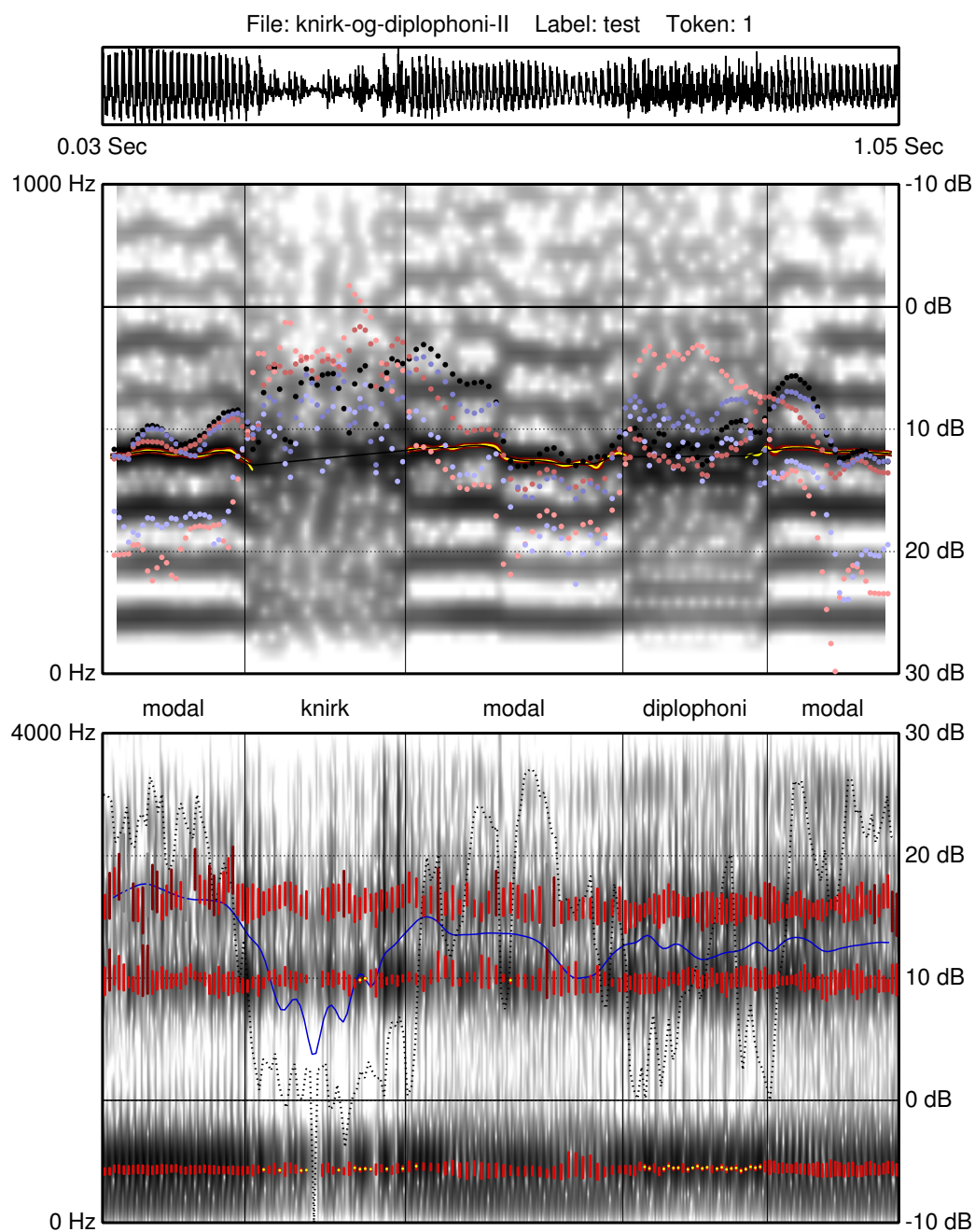
D.III. [i], f_0 -sweep. H1:H2 beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.



D.IV. [a], f_0 -sweep. H1:H2 beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.



D.V. Fortungerække. H1:H2 beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.

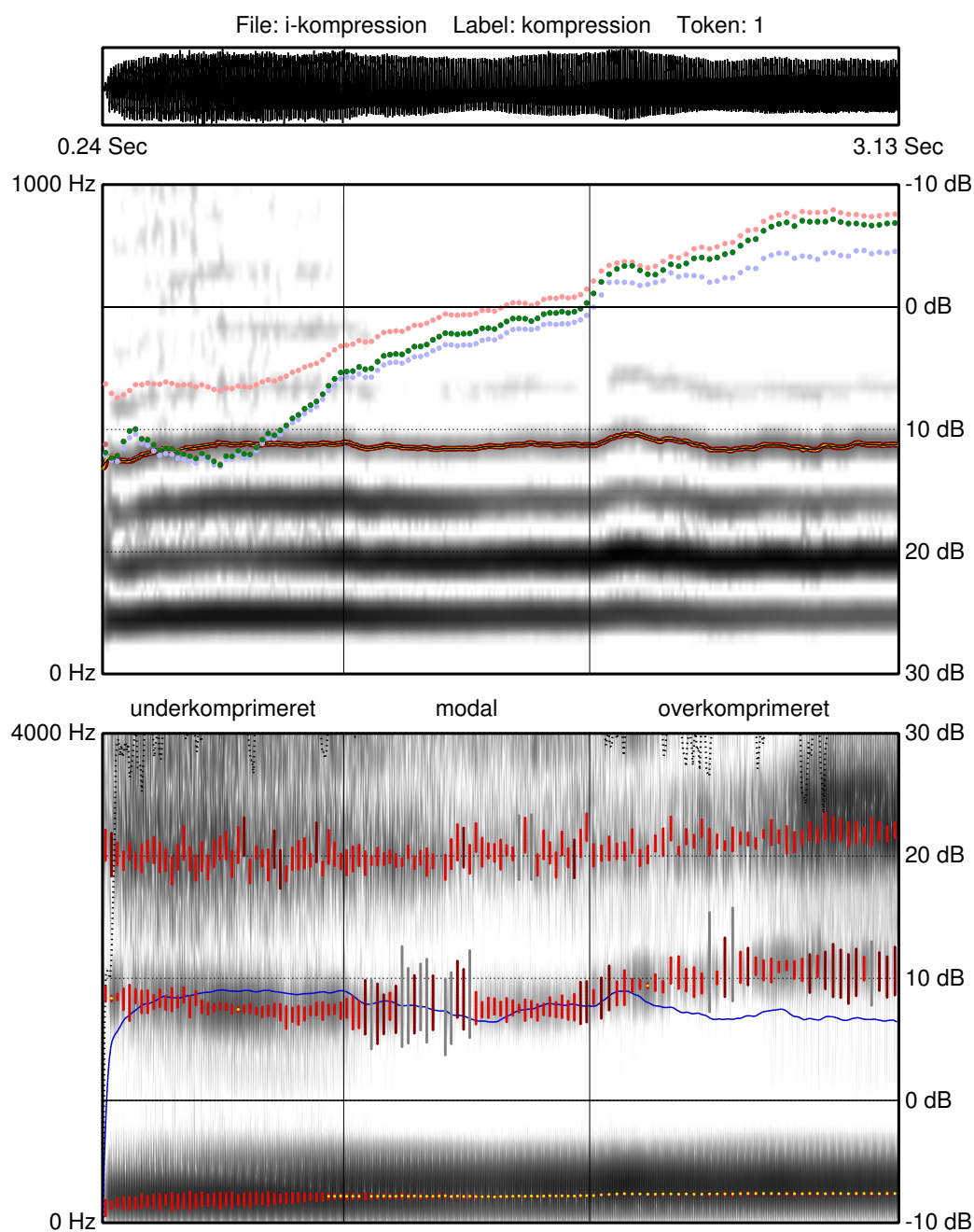


D.VI. [æ] med tre typer stemthed: modal, knirk og diplofoni.

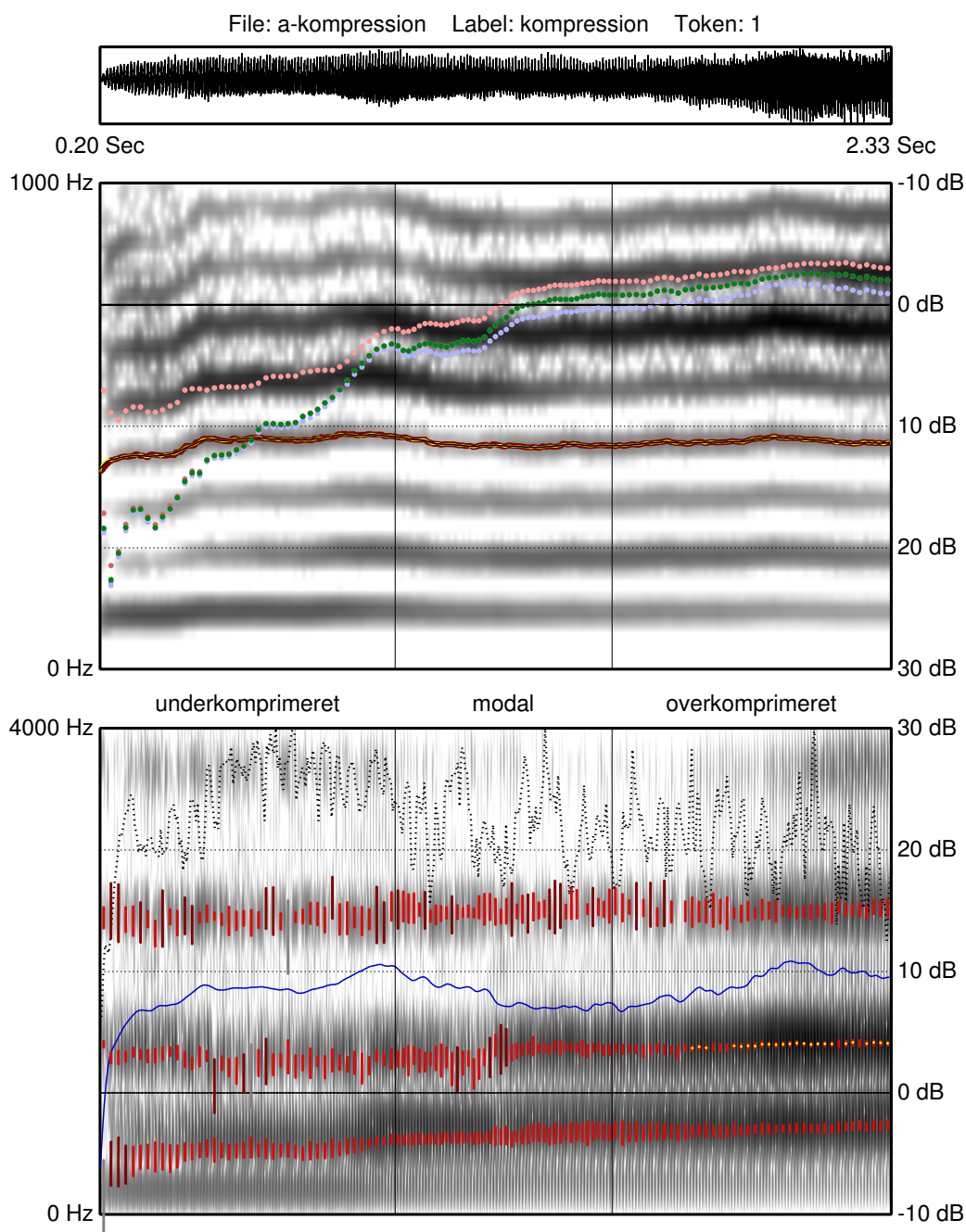
H1:H2 beregnet ud fra f_0 med offset af forskellig grad:

lys blålig: + 12.5 %, blålig: + 6.25 %, sort: 0 %, rødlig: - 6.25 %, lys rødlig: - 12.5 %.

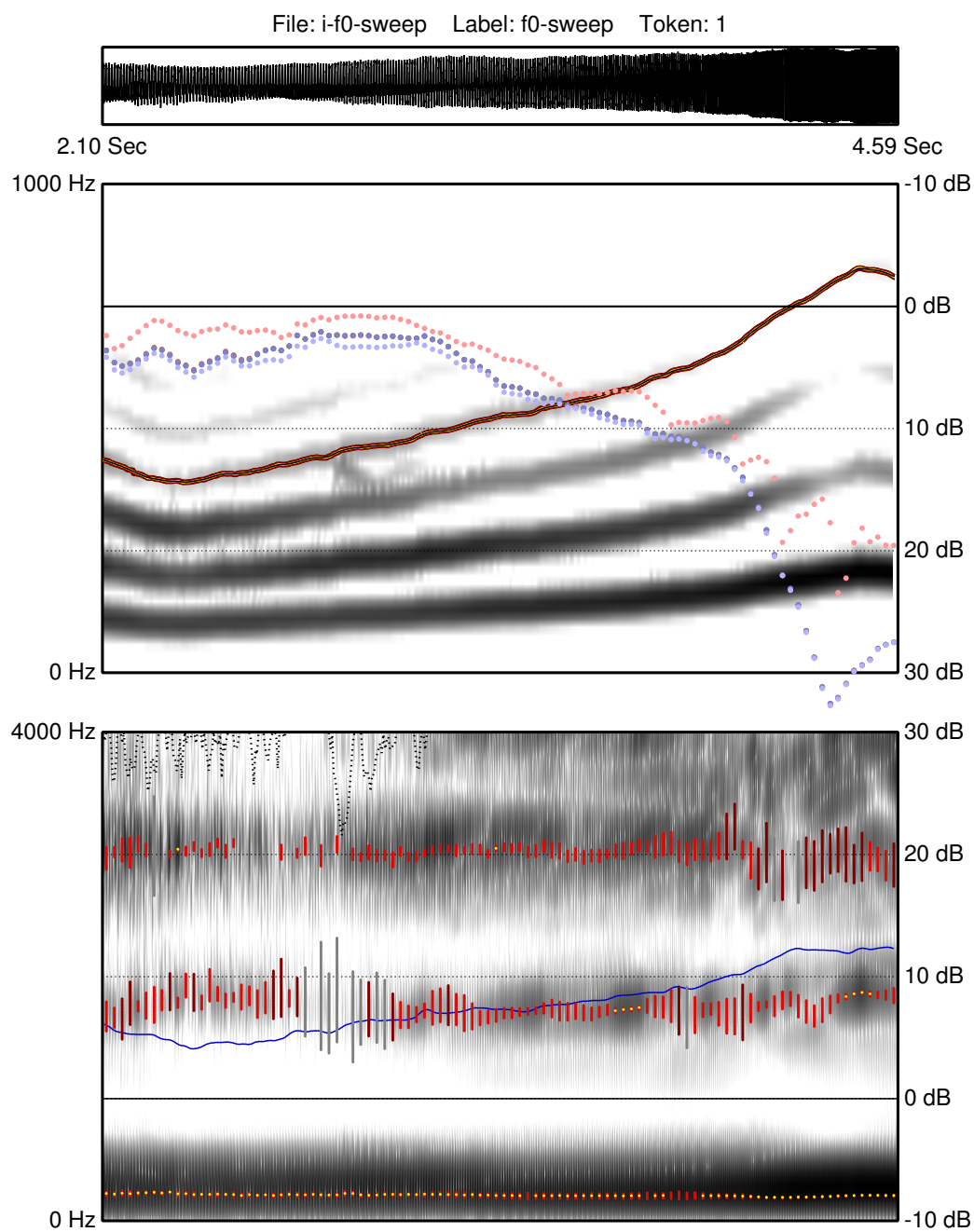
E Test af BED's følsomhed for fejl i f_0 -bestemmelsen



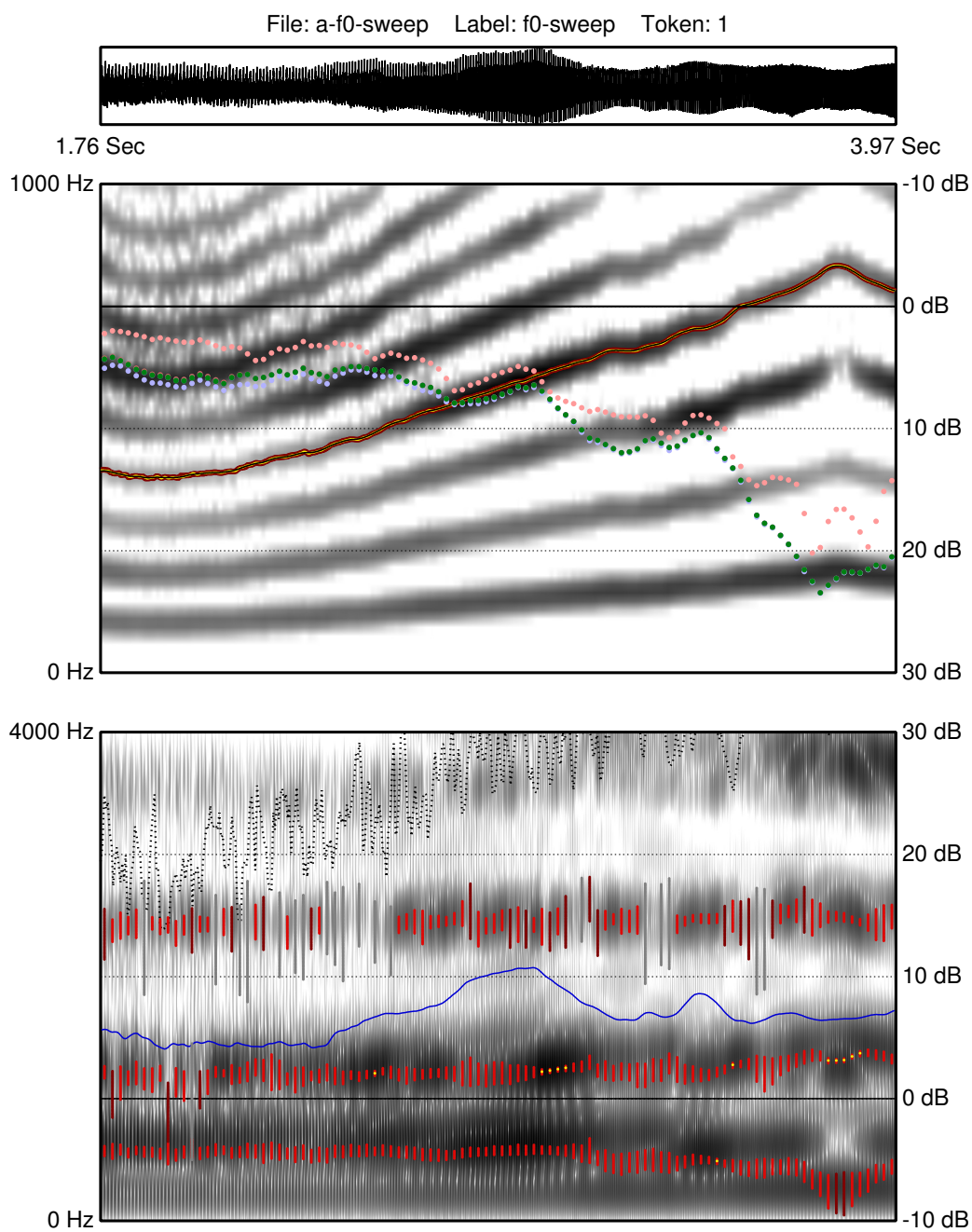
E.I. [i] med kompression varierende fra underkomprimeret til overkomprimeret.
 BED beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %.



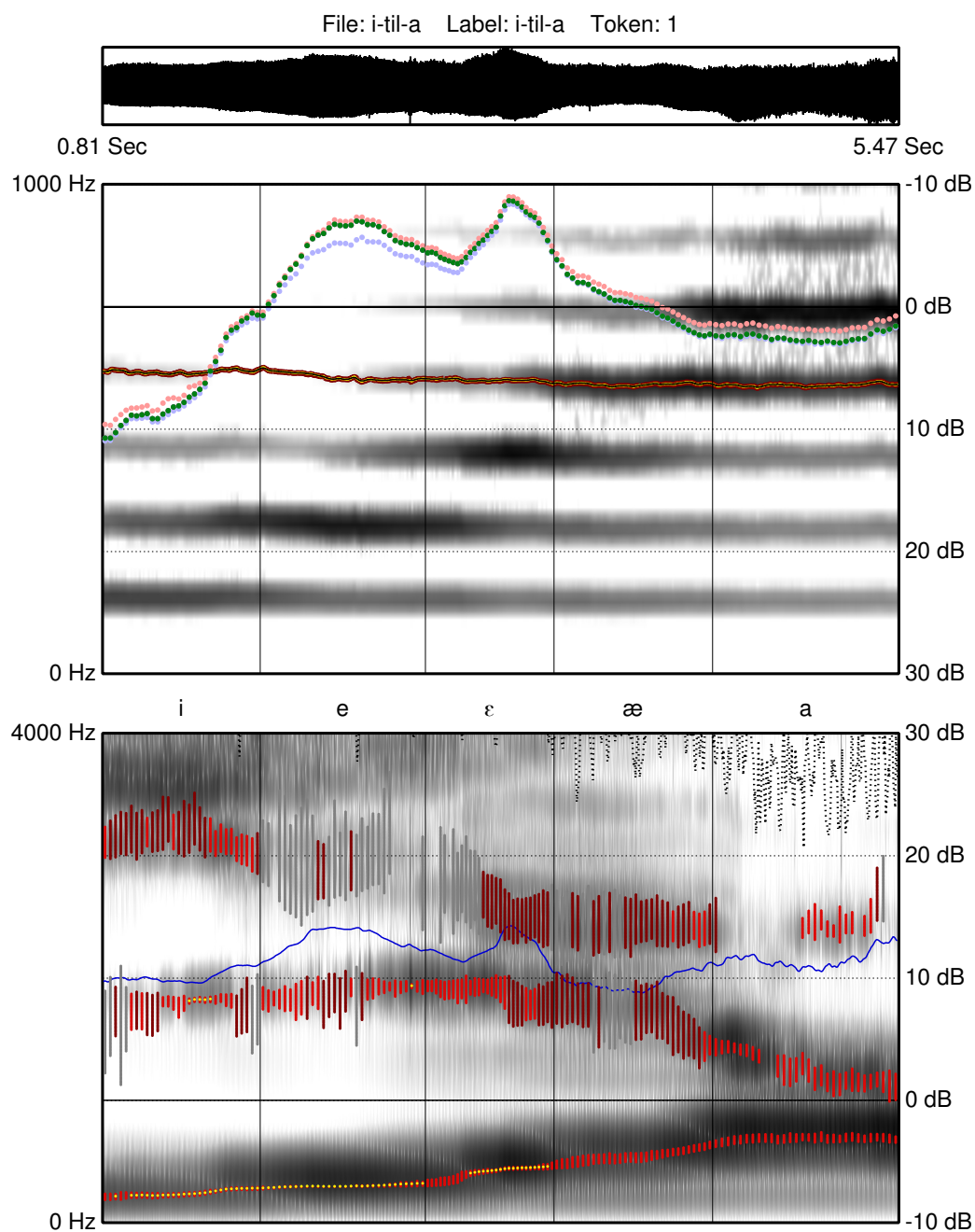
E.II. [a] med kompression varierende fra underkomprimeret til overkomprimeret.
 BED beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %.



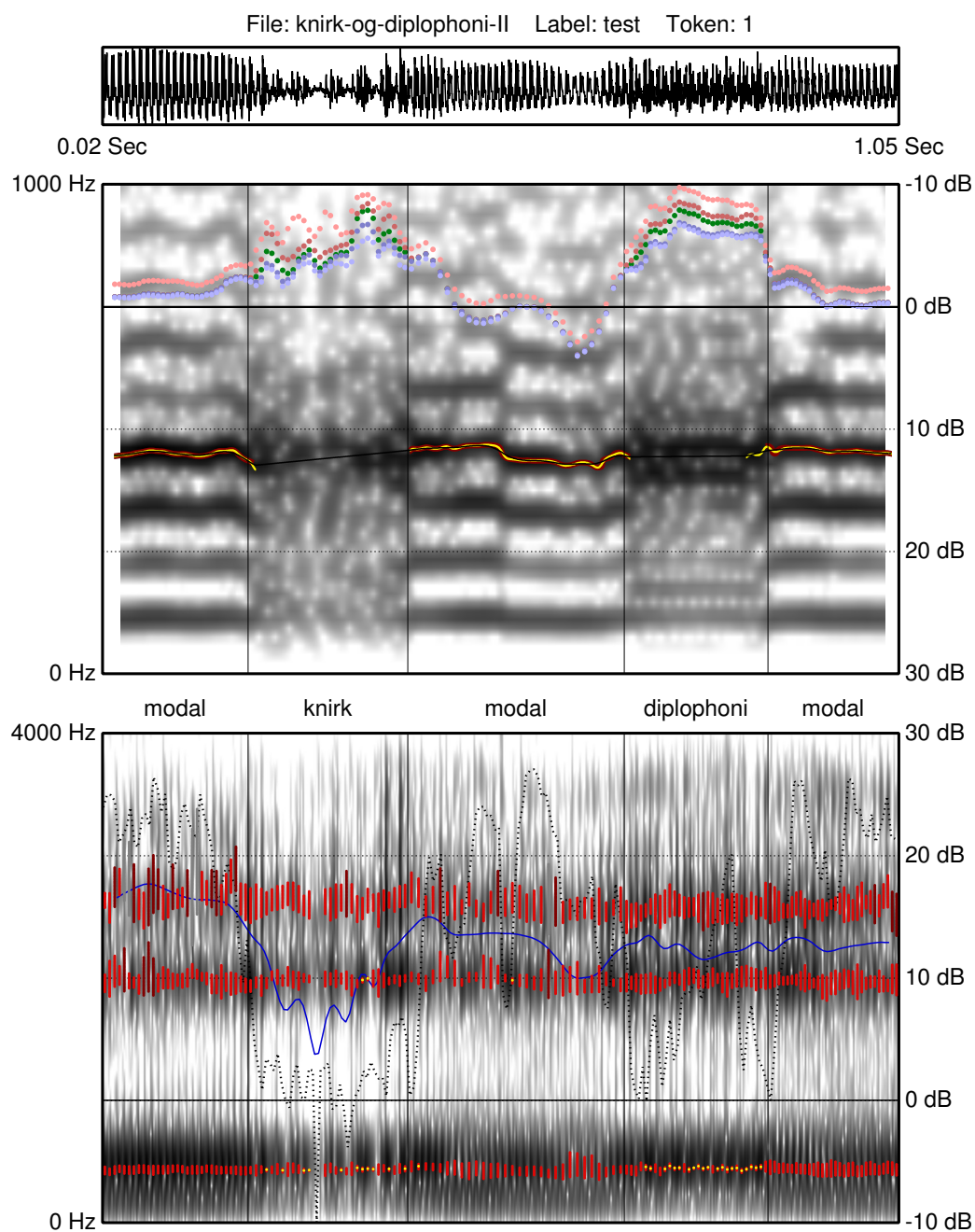
E.III. [i], f_0 -sweep. BED beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %.



E.IV. [a], f_0 -sweep. BED beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %.



E.V. Fortungerække. BED beregnet ud fra f_0 med offset af forskellig grad:
 lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %.

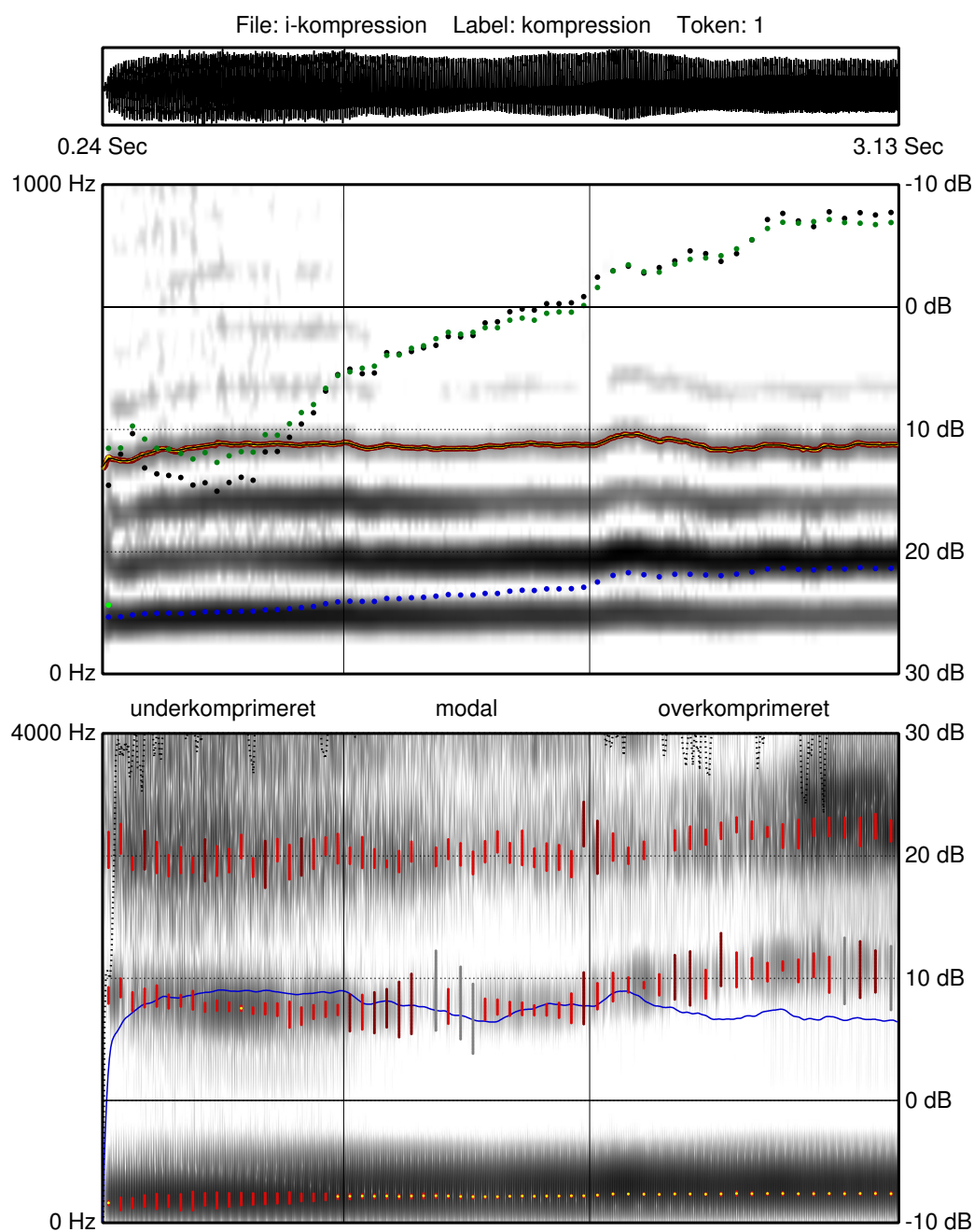


E.VI. [æ] med tre typer stemthed: modal, knirk og diplofoni.

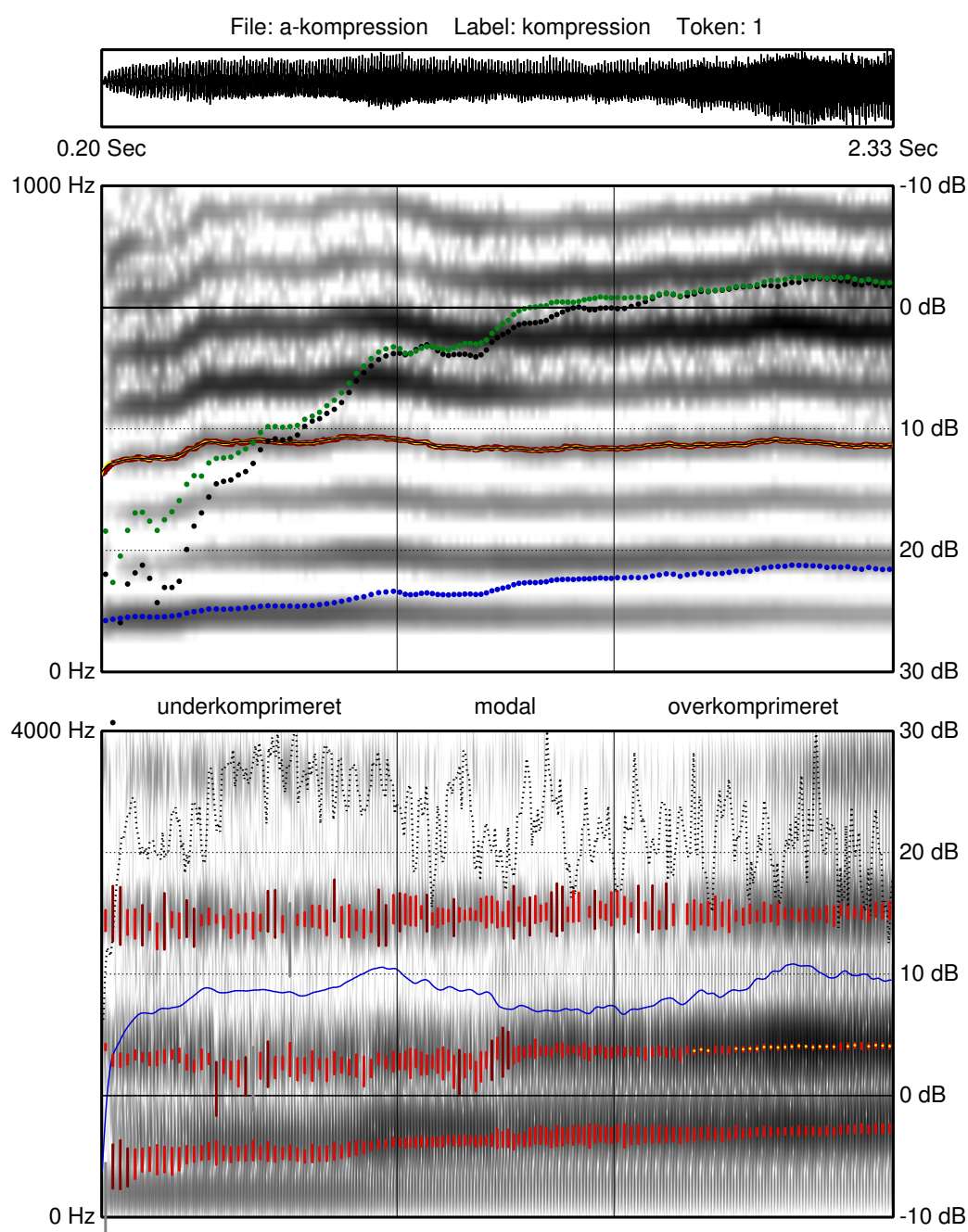
BED beregnet ud fra f_0 med offset af forskellig grad:

lys blålig: + 25 %, blålig: + 12.5 %, grøn: 0 %, rødlig: - 12.5 %, lys rødlig: - 25 %.

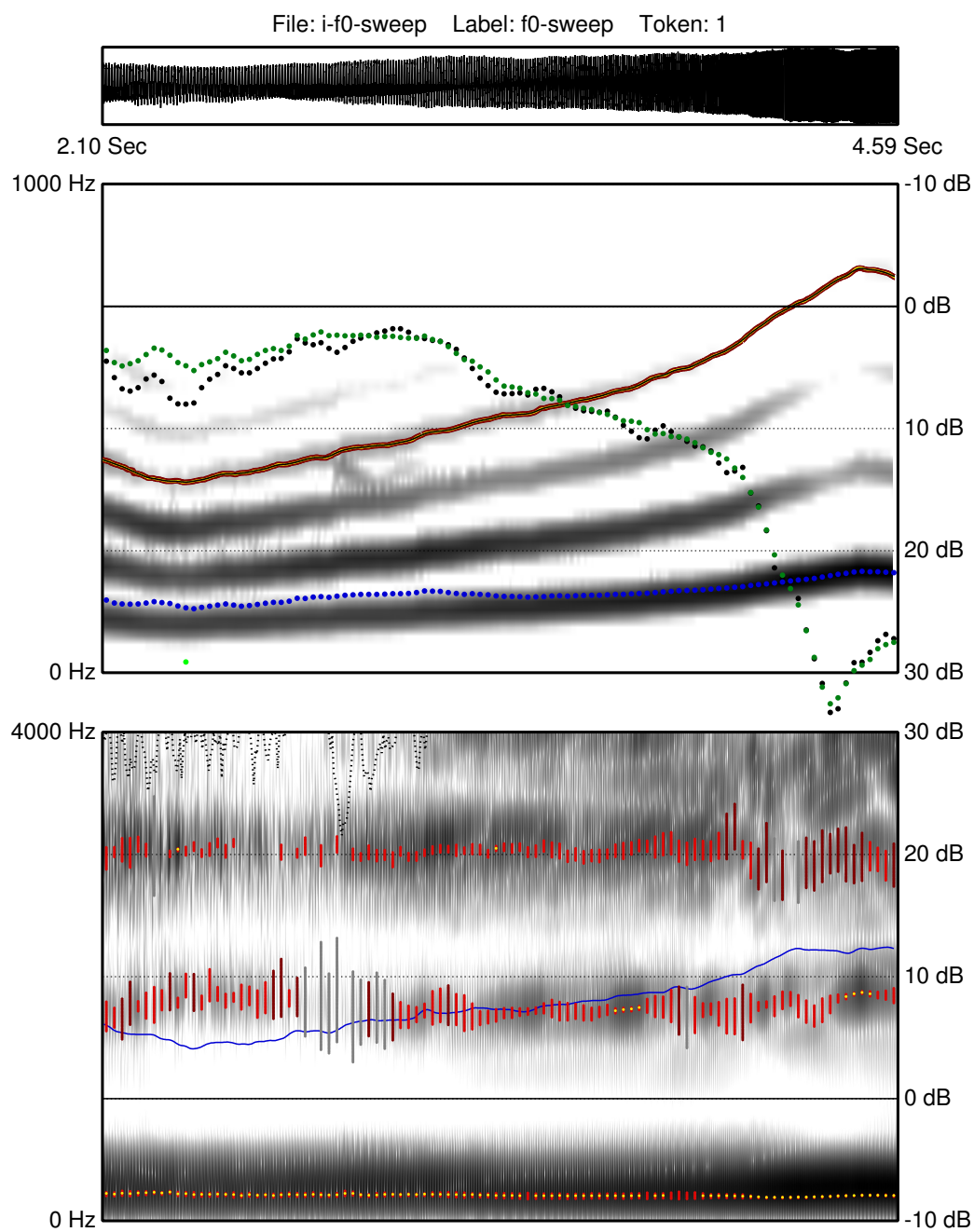
F Analyse (kurveblade) for testfiler



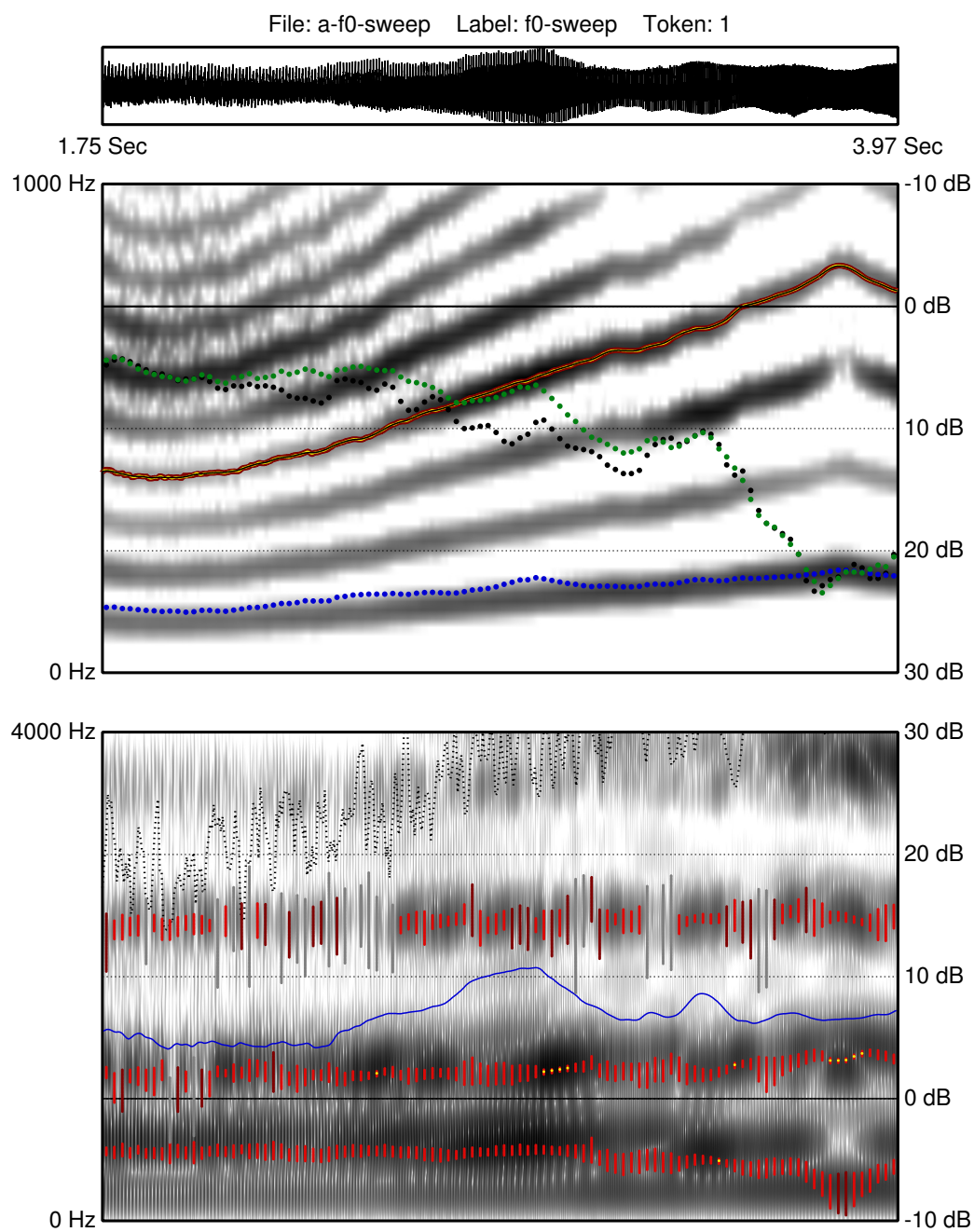
FI. [i], varierende kompression fra underkomprimeret til overkomprimeret.
Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå).



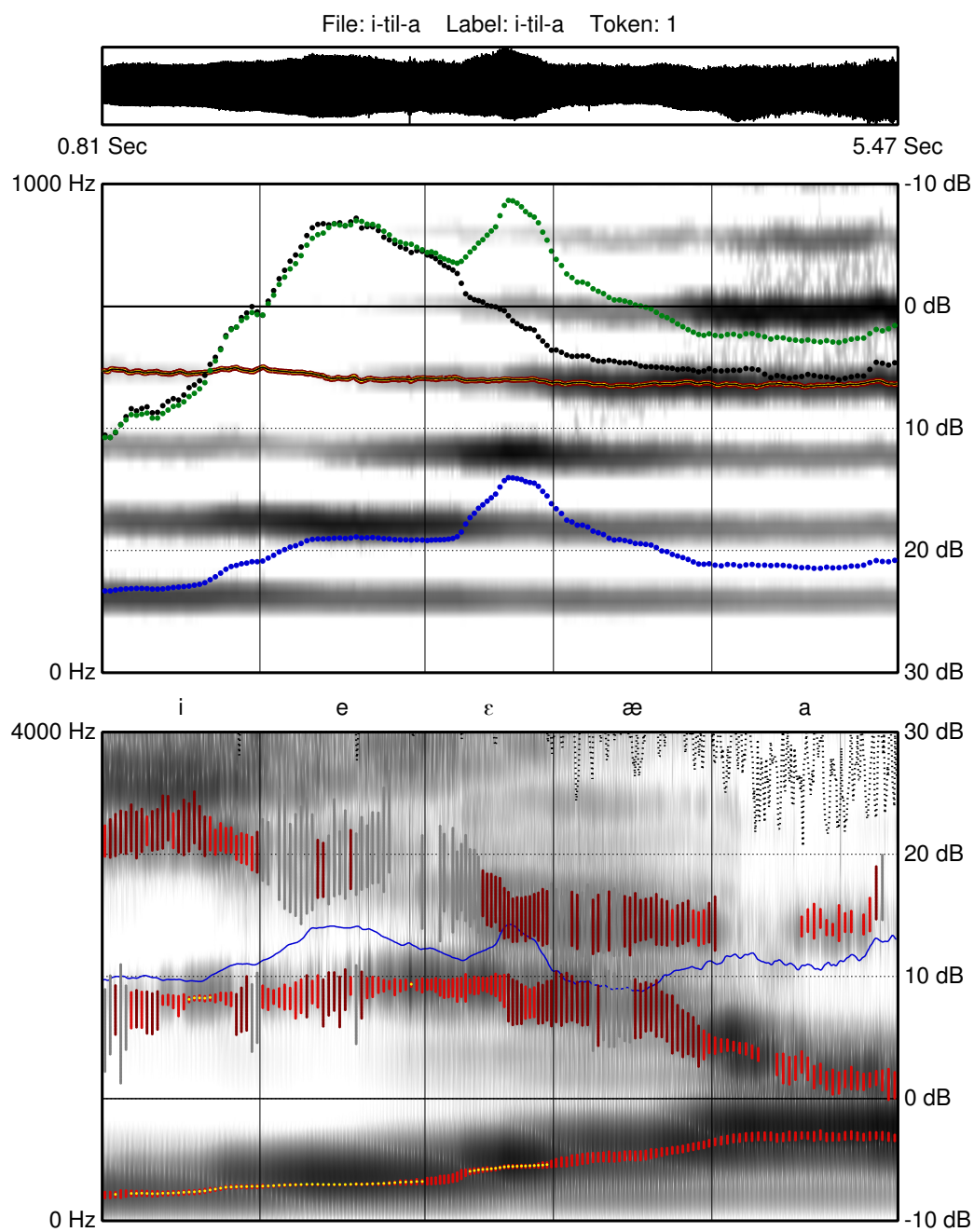
F.II. [a], varierende kompression fra underkomprimeret til overkomprimeret. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå).



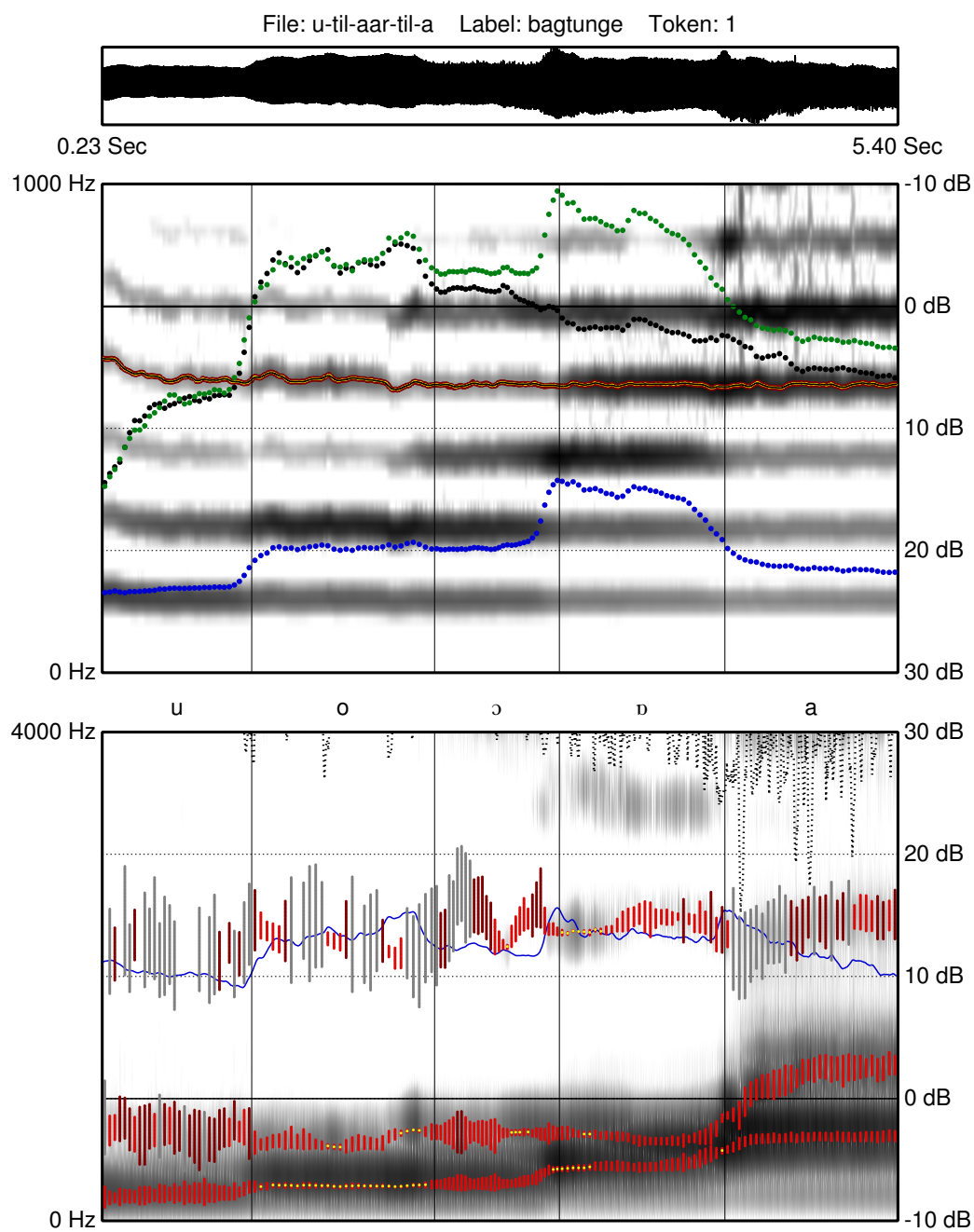
F.III. [i], varierende grundtone. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå). Den anslåede F1-værdi er markeret med en stiplet rød linje.



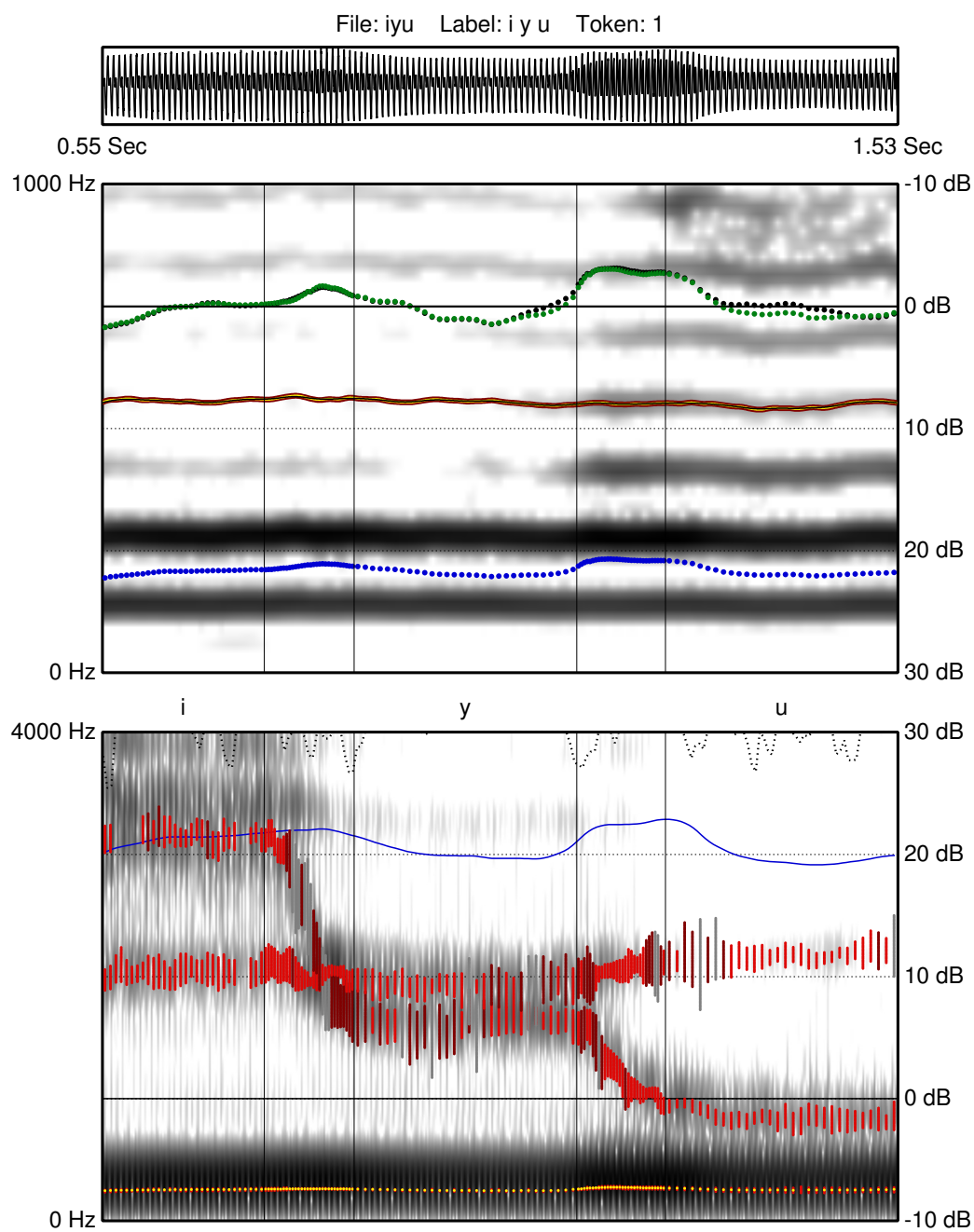
F.IV. [a], varierende grundtone. Sammenligning af kurveforløbene for H1:H2 (sort), BED (grøn) og CoG (blå). Den omtrentlige F1-værdi er markeret med en stipleet rød linje.



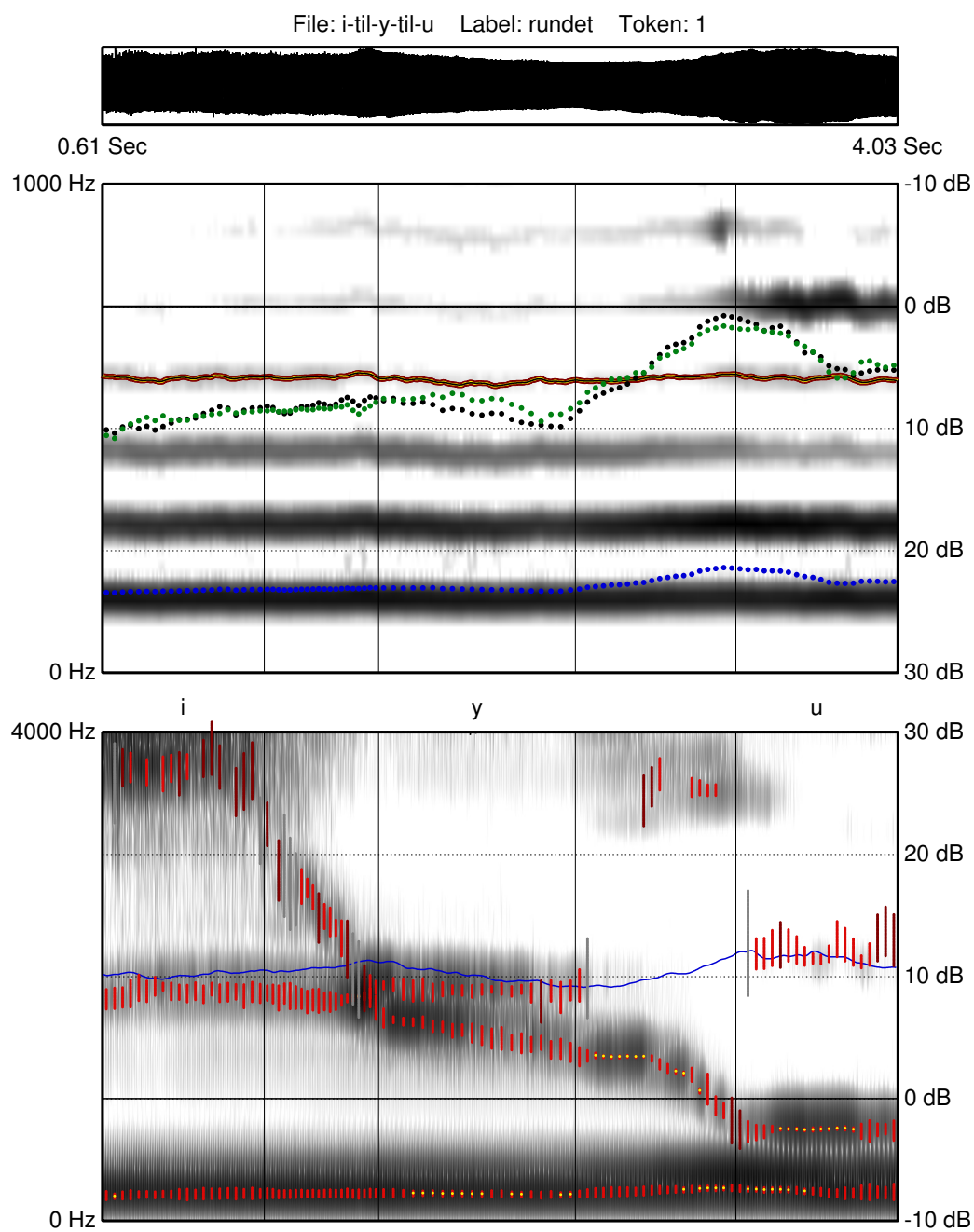
F.V. Serie af urundede fortunevokaler plus [a]. H1:H2 (sort), BED (grøn) og CoG (blå).



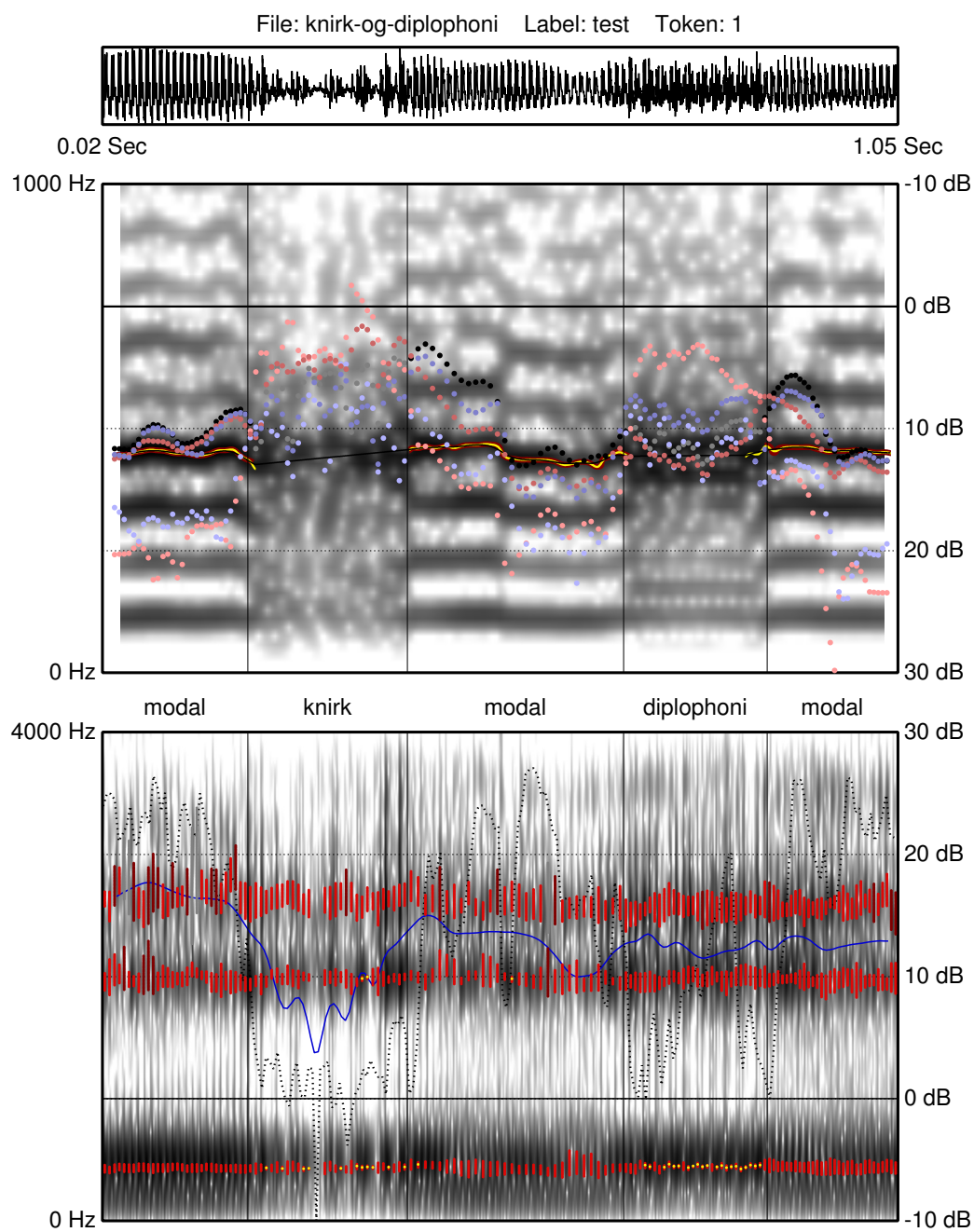
F.VI. Serie af bagtungevokaler plus [a]. H1:H2 (sort), BED (grøn) og CoG (blå).



F.VII.1 Serie af svævre vokaler: [i], [y] og [u]. H1:H2 (sort), BED (grøn) og CoG (blå).



F.VII.2 Serie af svævre vokaler: [i], [y] og [u]. H1:H2 (sort), BED (grøn) og CoG (blå). Alternativ optagelse.



F.VIII. Afprøvning af $H1:H1\frac{1}{2}$ -målet på optagelse af [æ] med tre typer stemthed: modal, knirk og diplofoni.

G Testord

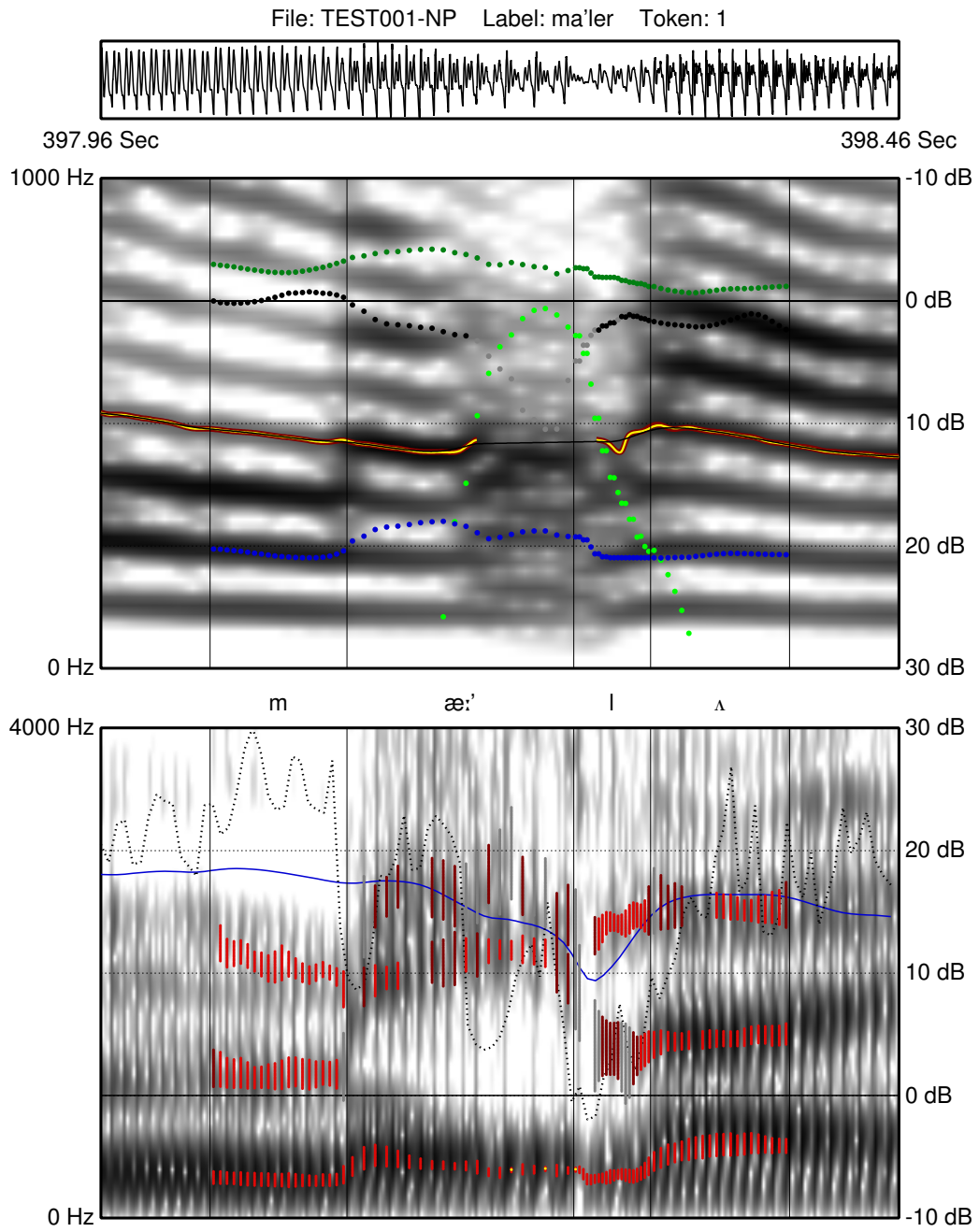
Oversigt over testordene med lydskrift plus sætningsmateriale. Enkelte ord forekommer i flere sætninger. Lydskrift i parentes repræsenterer mere distinkt men mindre sandsynlig udtale.

a'sen	[¹ æ:'sɐ] ([¹ æ:'sɛn])	Arne var et stædigt <u>asen</u> at arbejde sammen med.
aser	[¹ æ:sʌ]	Det er spændende at høre om <u>aser</u> og guder i den nordiske mytologi.
baller	[¹ bʌlʌ]	Bonden lod et par af de store <u>baller</u> blive på marken.
ball'er	[¹ bʌl'ʌ]	De to <u>baller</u> bliver arrangeret af eleverne.
barneben	[¹ bɑ:nə,be:'n]	Man kunne kun se et enkelt <u>barneben</u> stritte ud af halmen.
bar'net	[¹ bɑ:'nɛð]	Det vigtigt, at <u>barnet</u> for lov at lege.
barnlig	[¹ bɑ:nli]	En vis <u>barnlig</u> nysgerrighed skader ikke.
bar'nsben	[¹ bɑ:'ns,be:'n]	Arne havde lige fra <u>bar'nsben</u> drømt om at blive maler.
bar'sel	[¹ bɑ:'sɛ] ([¹ bɑ:'sɛl])	Den nye lærer skal på <u>barsel</u> efter sommerferien.
esser	[¹ ɛsʌ]	Man skulle tro, at læreren havde <u>esser</u> gemt i ærmet.
fa'gs	[¹ fæ:'s]	Det er vigtigt at alle <u>fags</u> undervisningslokaler er orden.
far's	[¹ fɑ:'s]	Husholdningslæreren havde købt <u>fars</u> til frikadeller.
fil'm	[¹ fi'l'm]	Det er meningen, at der skal vises <u>film</u> i den nye hal.
fil'm	[¹ fi'l'm]	Eleverne så en <u>film</u> om forskellige lande.
fi'n	[¹ fi:'n]	Det var en <u>fin</u> gammel skole, som lå ned til en eng.
fine	[¹ fi::n] ([¹ fi:nə])	Eleverne havde lavet nogle <u>fine</u> tegninger.
fir's	[¹ fi_ʌ's]	Arne vidste, at der skulle bruges <u>firs</u> liter maling.
fir's	[¹ fi_ʌ's]	Der blev brugt <u>firs</u> liter maling til skolens facade.
fi's	[¹ fi:'s]	Man kunne godt lave lidt <u>fis</u> i timerne.
hal'	[¹ hal']	Børnene leger i den nye <u>hal</u> med fodboldmålene.
hal'm	[¹ hal'm]	Bonden havde <u>halm</u> i laden.
kæmpeballer	[¹ kʰɛmpebʌlʌ]	Den nye mejetærsker laver <u>kæmpeballer</u> i stedet for små.
kæmpeball'er	[¹ kʰɛmpebʌl'ʌ]	Der bliver holdt to <u>kæmpeballer</u> på skolen.
læ'ner	[¹ lɛ:'nʌ]	Arne står og <u>læner</u> sig op ad stigen.
læser	[¹ lɛ:sʌ]	Ib var den hurtigste <u>læser</u> i klassen.
læ'ser	[¹ lɛ:'sʌ]	Lærlingen <u>læser</u> brugsanvisningerne grundigt.
læsser	[¹ lɛsʌ]	Lærlingen <u>læsser</u> varevognen med de tomme malerbøtter.
lammet	[¹ lʌmɛð]	Læreren var <u>lammet</u> af forundring over stilheden.
lamm'et	[¹ lʌm'ɛð]	Fåret og <u>lamm'et</u> gik på engen og græssede.
lar'm	[¹ lɑ:'m]	Der var mere <u>larm</u> i klassen om fredagen.
lar'me	[¹ lɑ::m] ([¹ lɑ:mə])	Lad være med at <u>lar'me</u> så meget som, i plejer.
lar'men	[¹ lɑ:ɱ] ([¹ lɑ:'mɛn])	Man kunne høre <u>larmen</u> på lang afstand.
lar'met	[¹ lɑ:mɛð]	Der blev råbt og <u>lar'met</u> i klasseværelset.

li'men	[ˈli:m̥] ([ˈli:m̥ən])	Drengene havde gemt <u>limen</u> i skuffen.
limet	[ˈli:m̥øð]	Eleverne havde klippet og <u>limet</u> i formningstimen.
ma'l	[mæ:l]	Tegn og <u>mal</u> som i vil.
male	[mæ:l] ([ˈmæ:lə])	Lærlingen skal <u>male</u> væggene igen.
maler	[ˈmæ:lʌ]	Arne er <u>maler</u> i Viby.
ma'ler	[ˈmæ:lʌ]	Lærlingen <u>maler</u> alle væggene.
mal'm	[mal'm]	Læreren forklarer hvad <u>malm</u> kan bruges til.
ma'ser	[ˈmæ:sʌ]	Drengene <u>maser</u> med halmballerne hele eftermiddagen.
mases	[ˈmæ:səs]	Halmballerne skal <u>mases</u> for at de kan være på vognen.
overma'l	[ˈbʷʌmæ:l]	Afrens og <u>overmal</u> graffitien samme dag.
overma'le	[ˈbʷʌmæ:l]	Lærlingen skulle <u>overmale</u> graffitien i dag.
overmaler	[ˈbʷʌmæ:lʌ]	Arne blev udnævnt til <u>overmaler</u> i firmaet.
overma'ler	[ˈbʷʌmæ:lʌ]	Lærlingen <u>overmaler</u> graffitien.
smi'l	[smi:l]	Læreren vender tilbage med et <u>smil</u> på læben.
smile	[smi:l] ([ˈsmi:lə])	Det er vigtigt at <u>smile</u> til eleverne.
smi'ler	[ˈsmi:lʌ]	Skoleinspektøren <u>smiler</u> til eleverne.
smilet	[ˈsmi:løð]	Der blev altid <u>smilet</u> en del på den skole.
taber	[ˈtʰæ:bʌ]	Det er aldrig sjovt, når nogen i skolegården råber <u>taber</u> efter én.
ta'ber	[ˈtʰæ:bʌ]	Eleverne tror, at læreren <u>taber</u> sit væddemål med maleren.
ta'g	[tʰæ:]	Der skulle både skiftes <u>tag</u> og males inden sommerferien.
tage	[tʰæ:]	Der er ved at blive lagt nye <u>tage</u> på alle skolens bygninger.
ti'	[tʰi:]	Der er kun plads til <u>ti</u> elever i skolekøkkenet.
tie	[tʰi:]	Læreren siger, at eleverne skal <u>tie</u> og høre efter.
udtaler	[ˈuðtʰæ:lʌ]	Elevernes <u>udtaler</u> af ordene var fejlfri.
udta'ler	[ˈuðtʰæ:lʌ]	Skoleinspektøren <u>udtaler</u> sig til lokalavisen.
udta'ler	[ˈuðtʰæ:lʌ]	Eleverne <u>udtaler</u> ordene, som de har lært.
varsel	[ˈvɑ:sɕ] ([ˈvɑ:səl])	Helt uden <u>varsel</u> begyndte det at regne.
vasen	[ˈvæ:sɕ] ([ˈvæ:sən])	Nøglerne lå ved siden af <u>vasen</u> med blomster.
æn'der	[ˈɛn'ʌ]	Der er to slags <u>ænder</u> i dammen.
æn'der	[ˈɛn'ʌ]	Der er både <u>ænder</u> og gæs i dammen ude på engen.
æn'der	[ˈɛn'ʌ]	Der går to <u>ænder</u> ude i skolegården.
æn'der	[ˈɛn'ʌ]	De to <u>ænder</u> tiltrækker sig meget opmærksomhed.
æn'der	[ˈɛn'ʌ]	Der plejer at være flere <u>ænder</u> om efteråret.
æn'der	[ˈɛn'ʌ]	Om vinteren er der næsten ingen <u>ænder</u> .
- fyld	-	Arne spiller kort med læreren i frokostpausen.

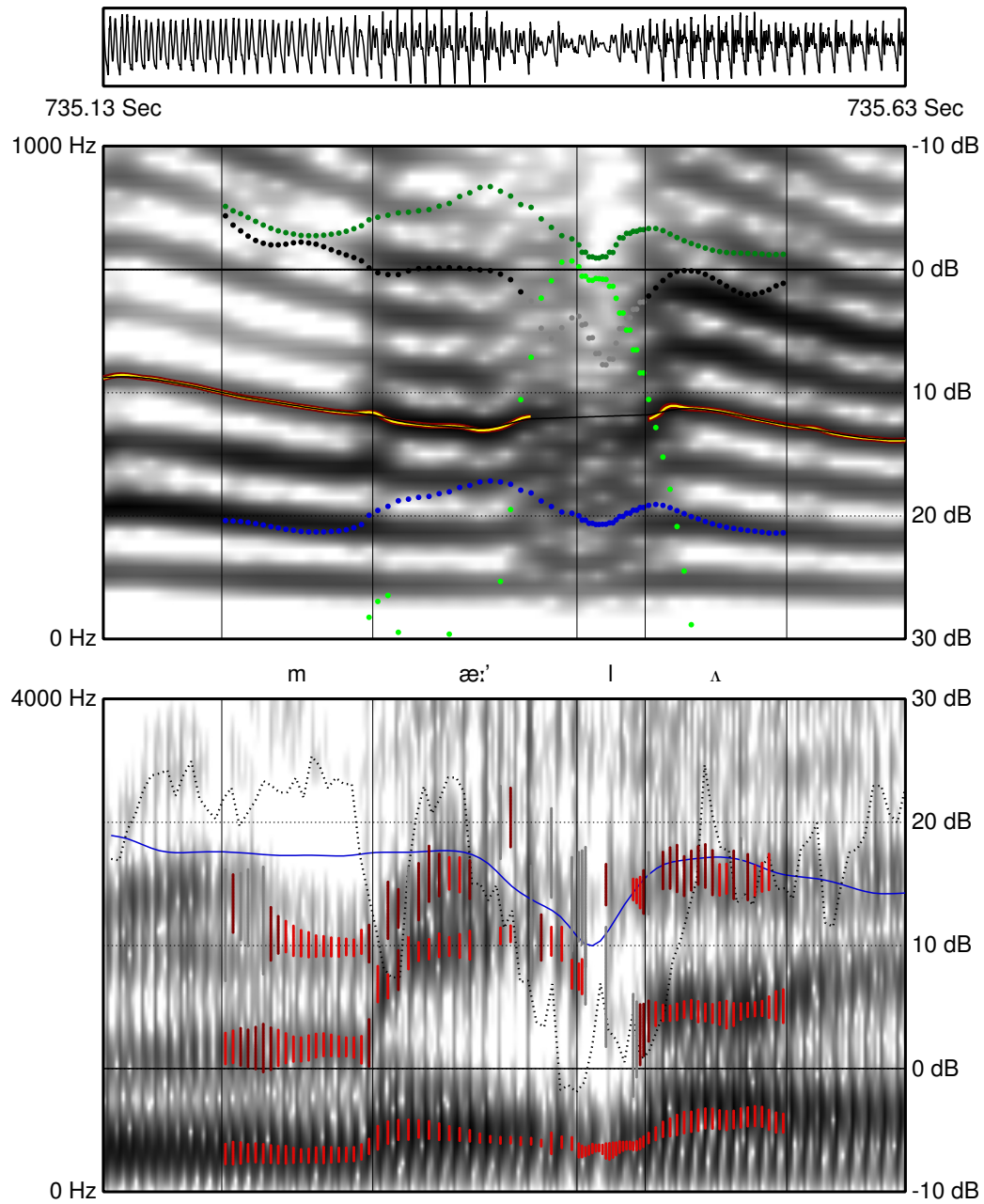
H Kurveblade for udvalgte testord

H.I ma'ler

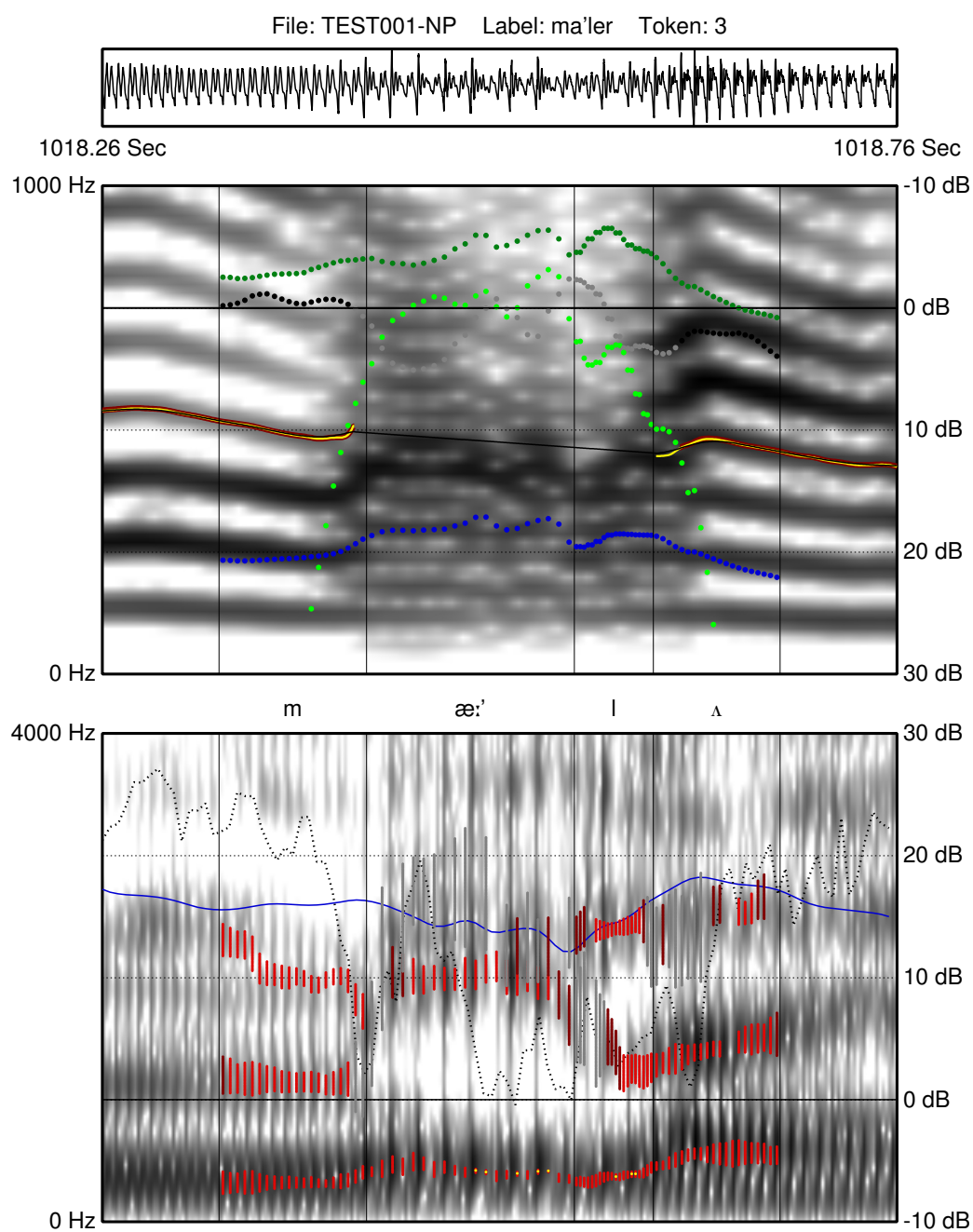


H.I.1. ma'ler token 1

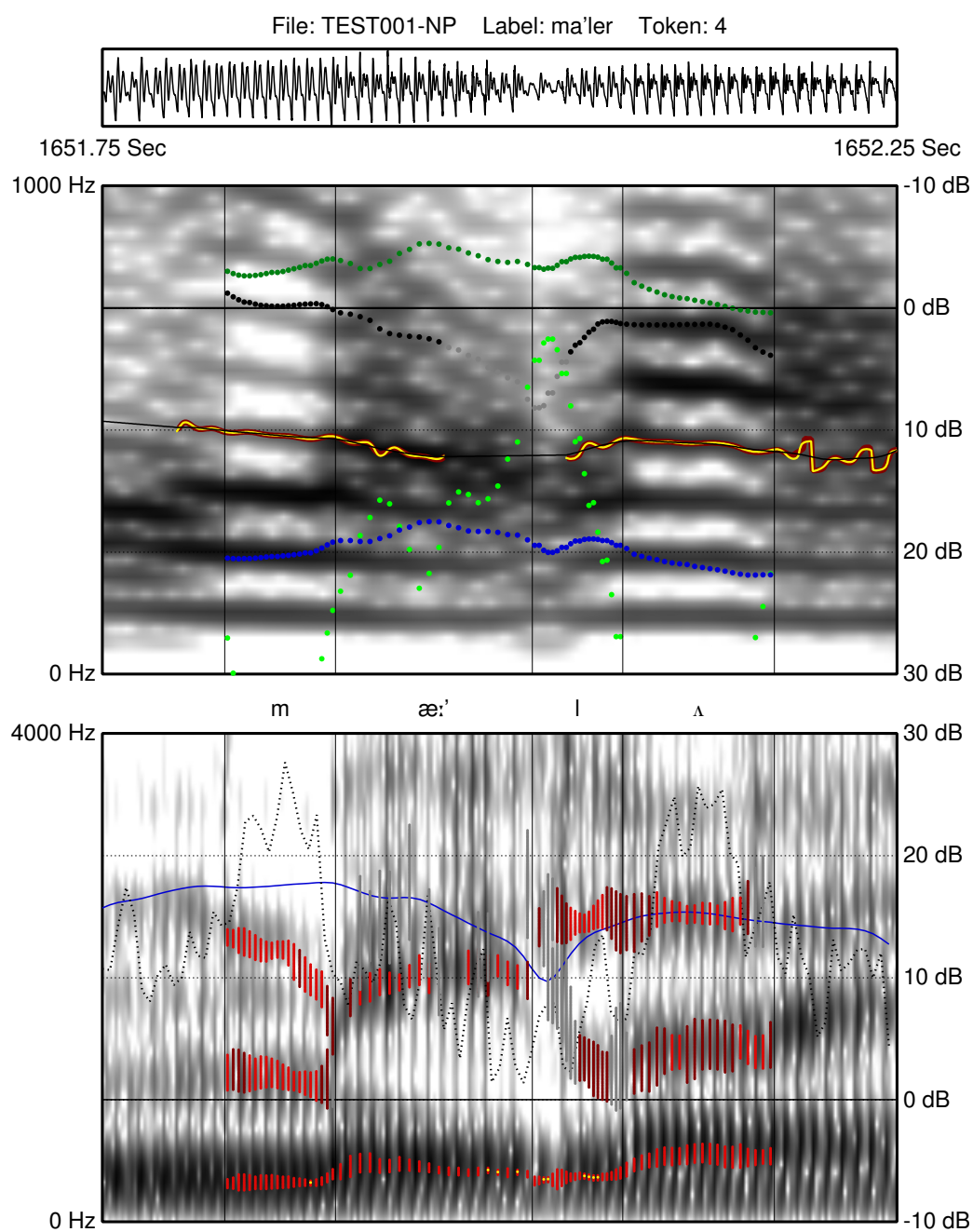
File: TEST001-NP Label: ma'ler Token: 2



H.I.2. ma'ler token 2

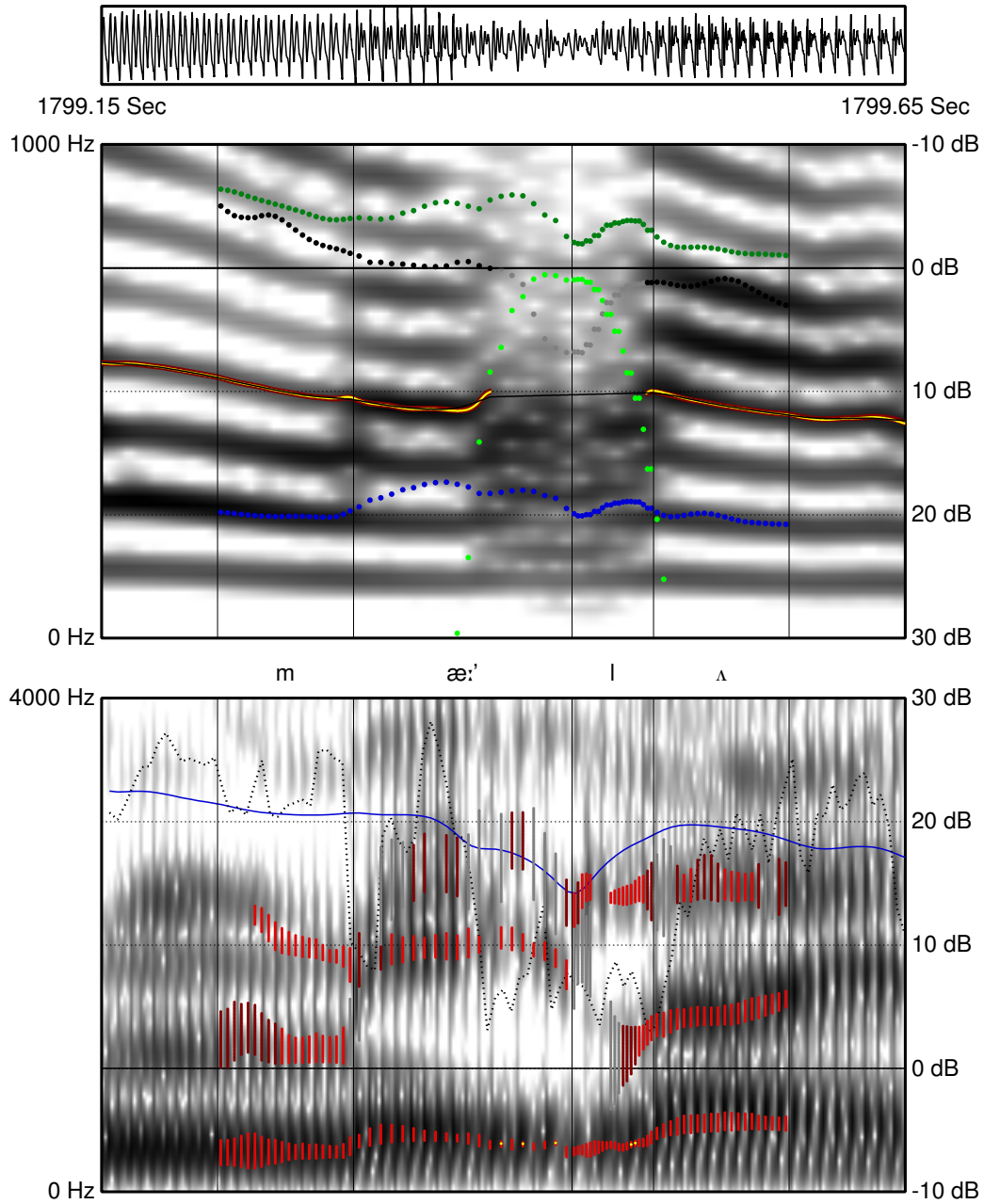


H.I.3. ma'ler token 3



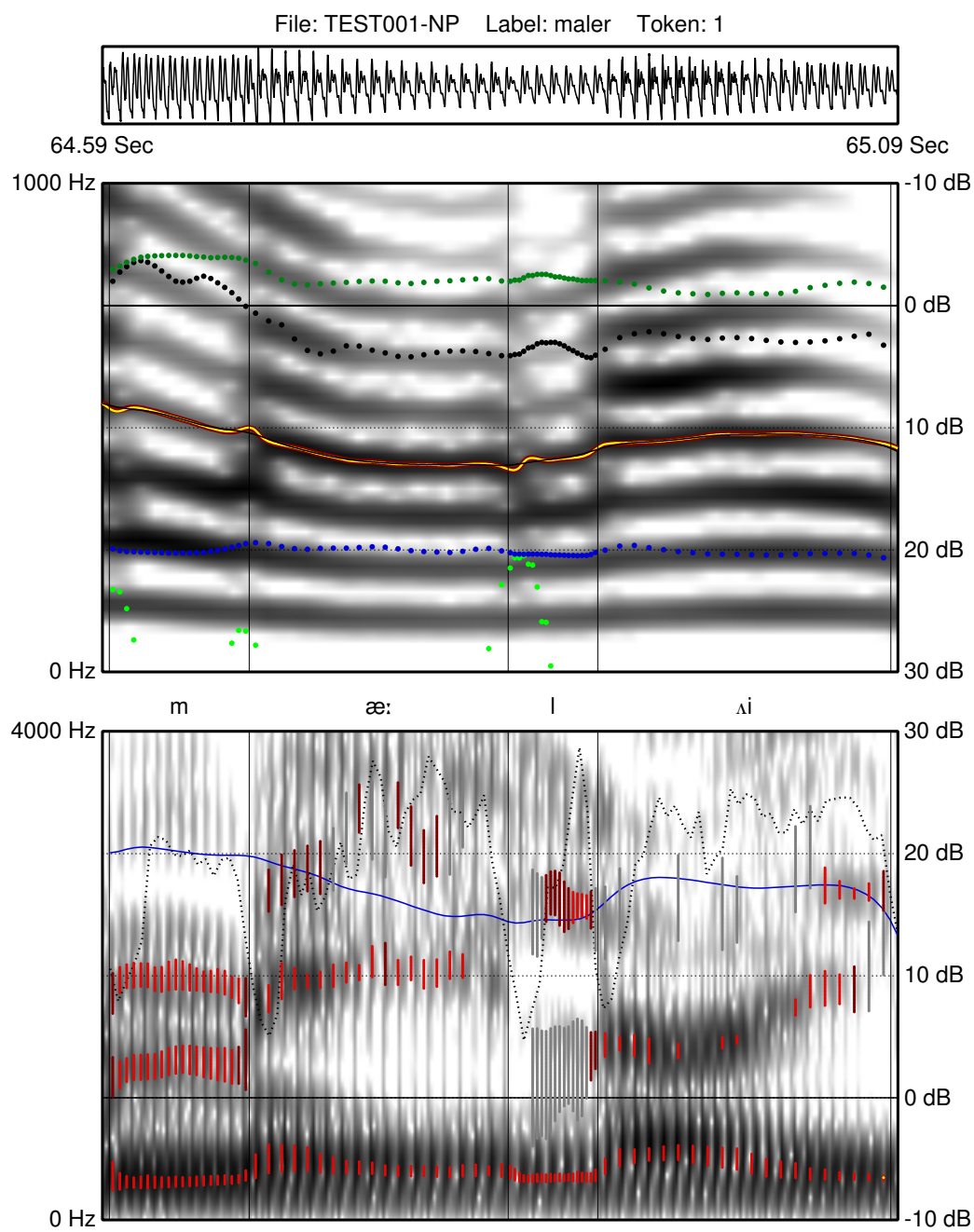
H.I.4. ma'ler token 4

File: TEST001-NP Label: ma'ler Token: 5



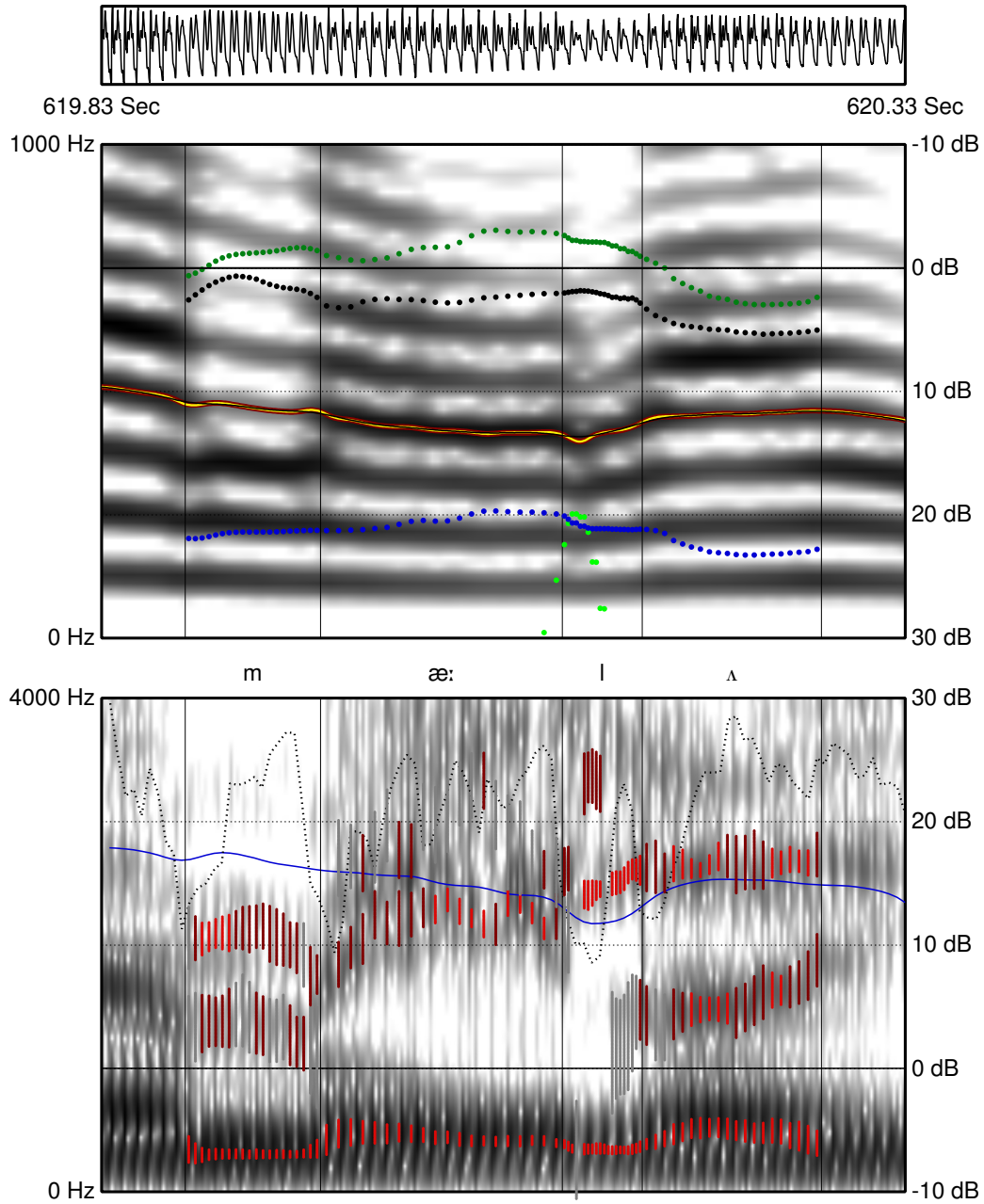
H.I.5. ma'ler token 5

H.II maler



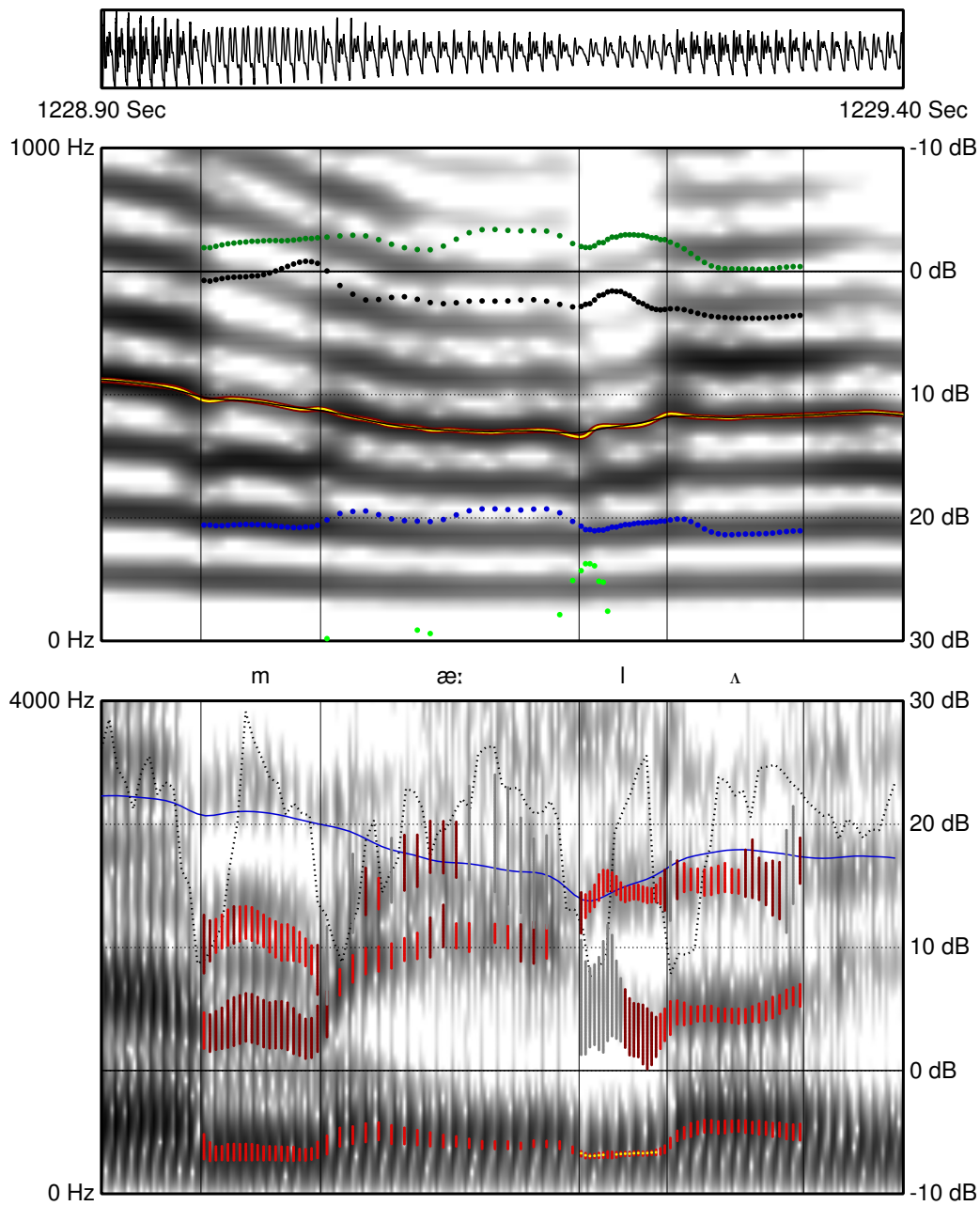
H.II.1. maler token 1

File: TEST001-NP Label: maler Token: 2



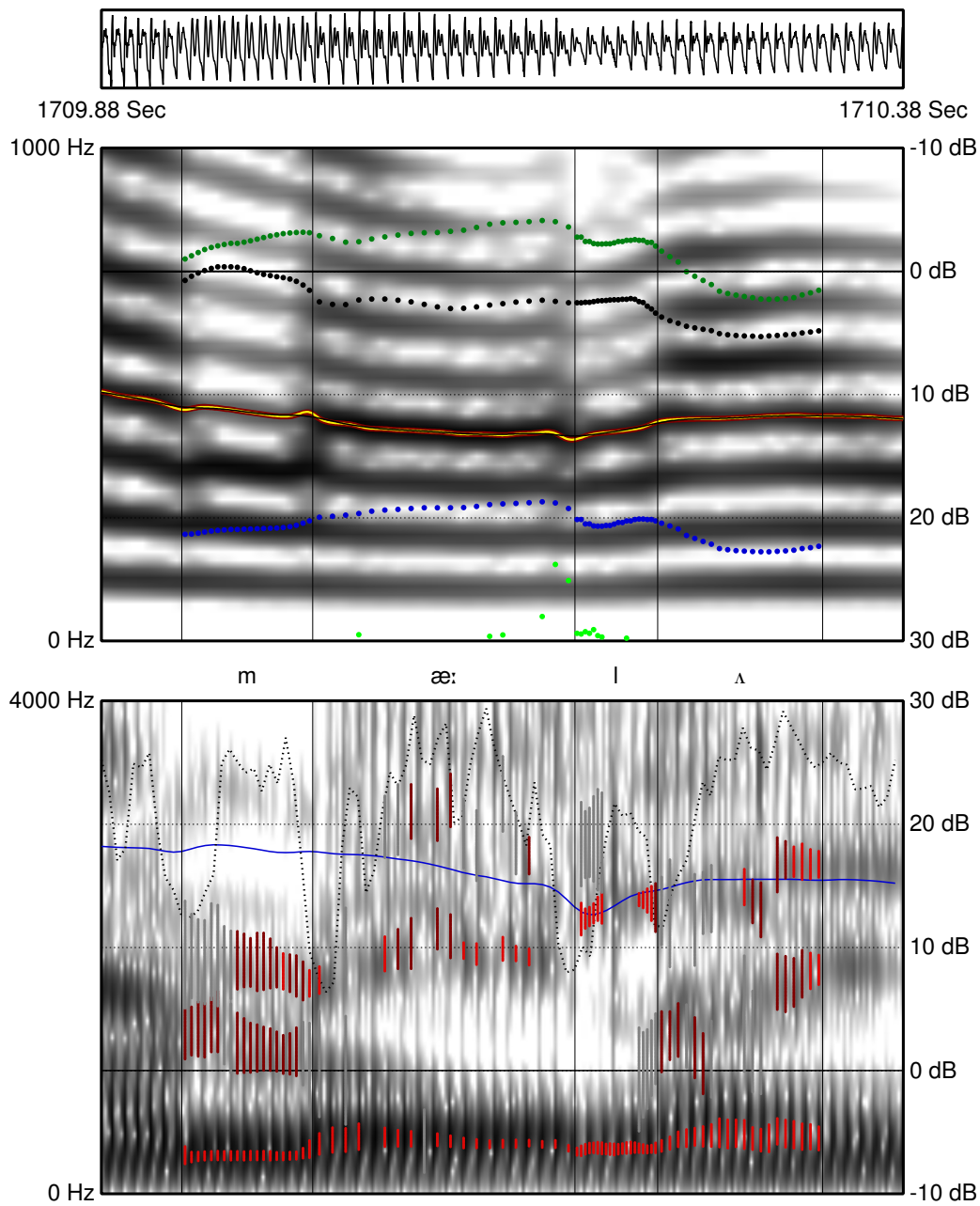
H.II.2. maler token 2

File: TEST001-NP Label: maler Token: 3



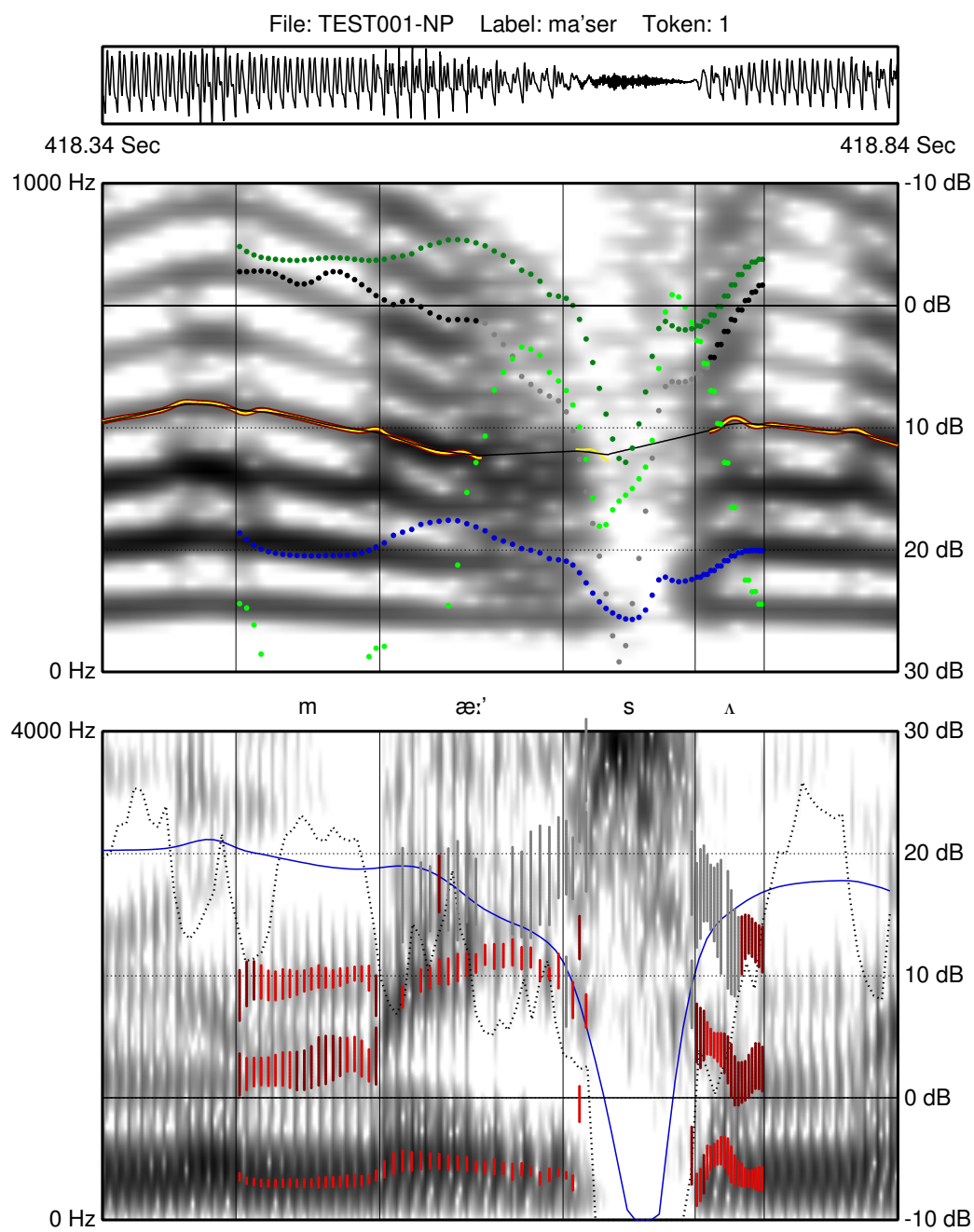
H.II.3. maler token 3

File: TEST001-NP Label: maler Token: 4



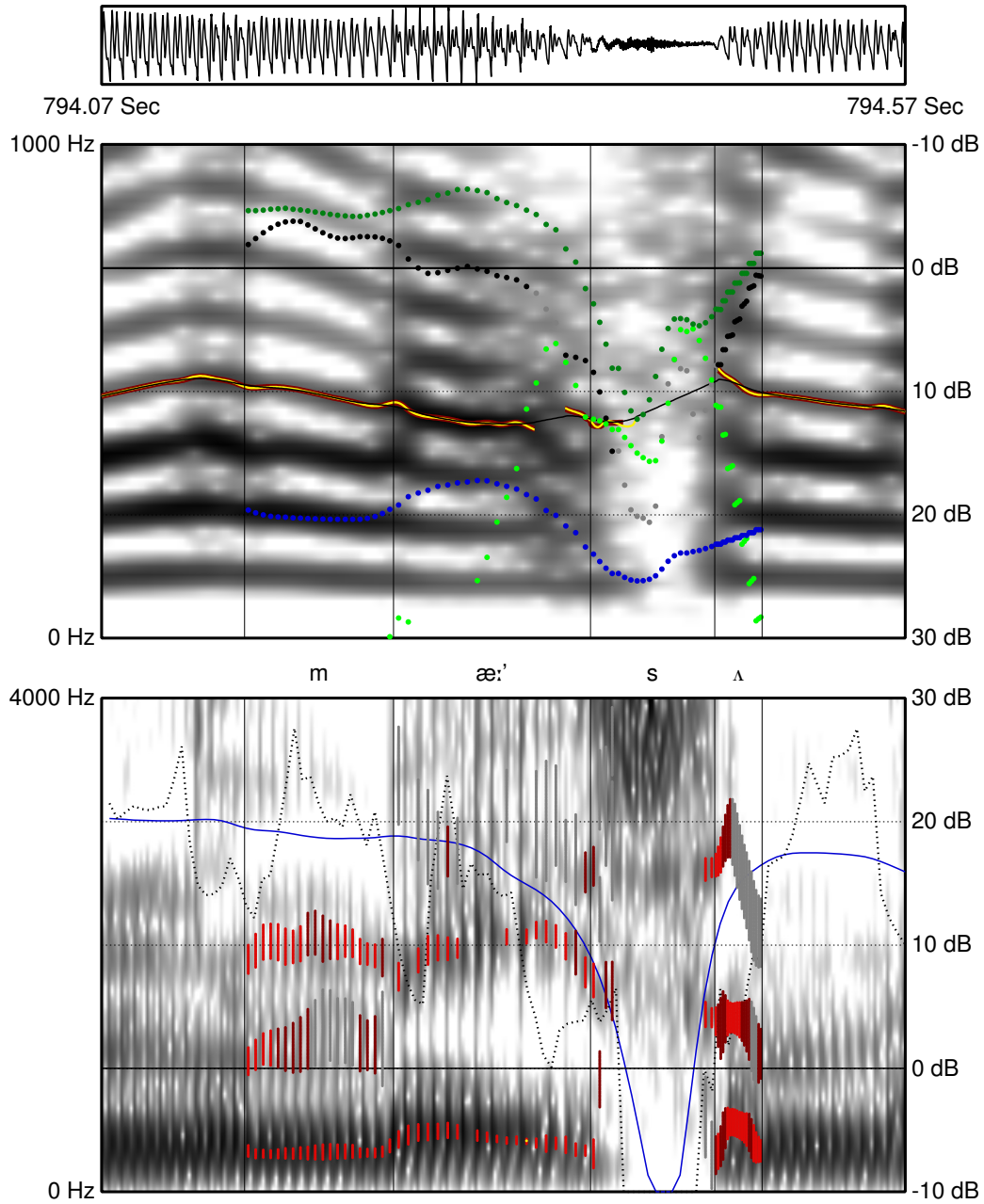
H.II.4. maler token 4

H.III ma'ser

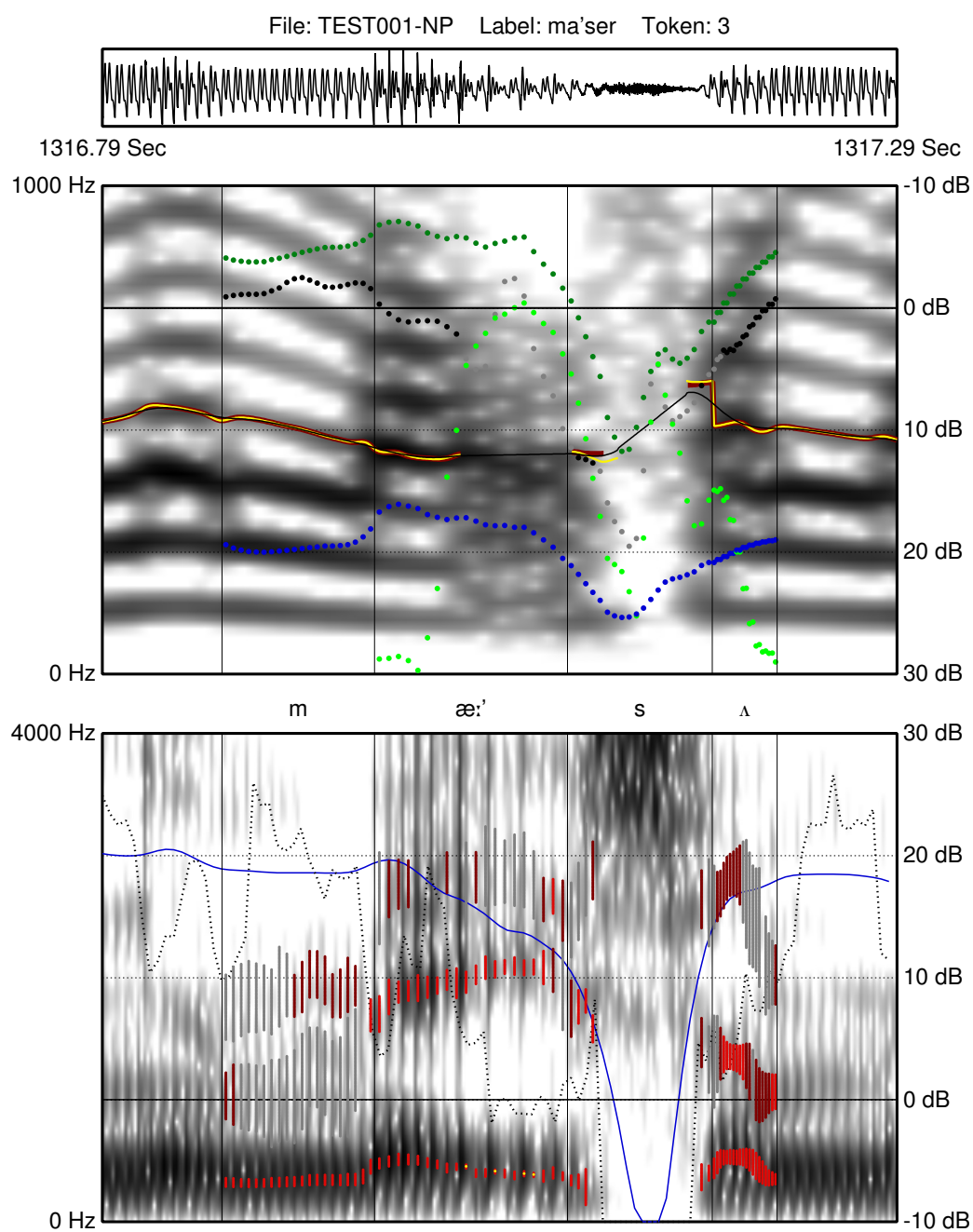


H.III.1. ma'ser token 1

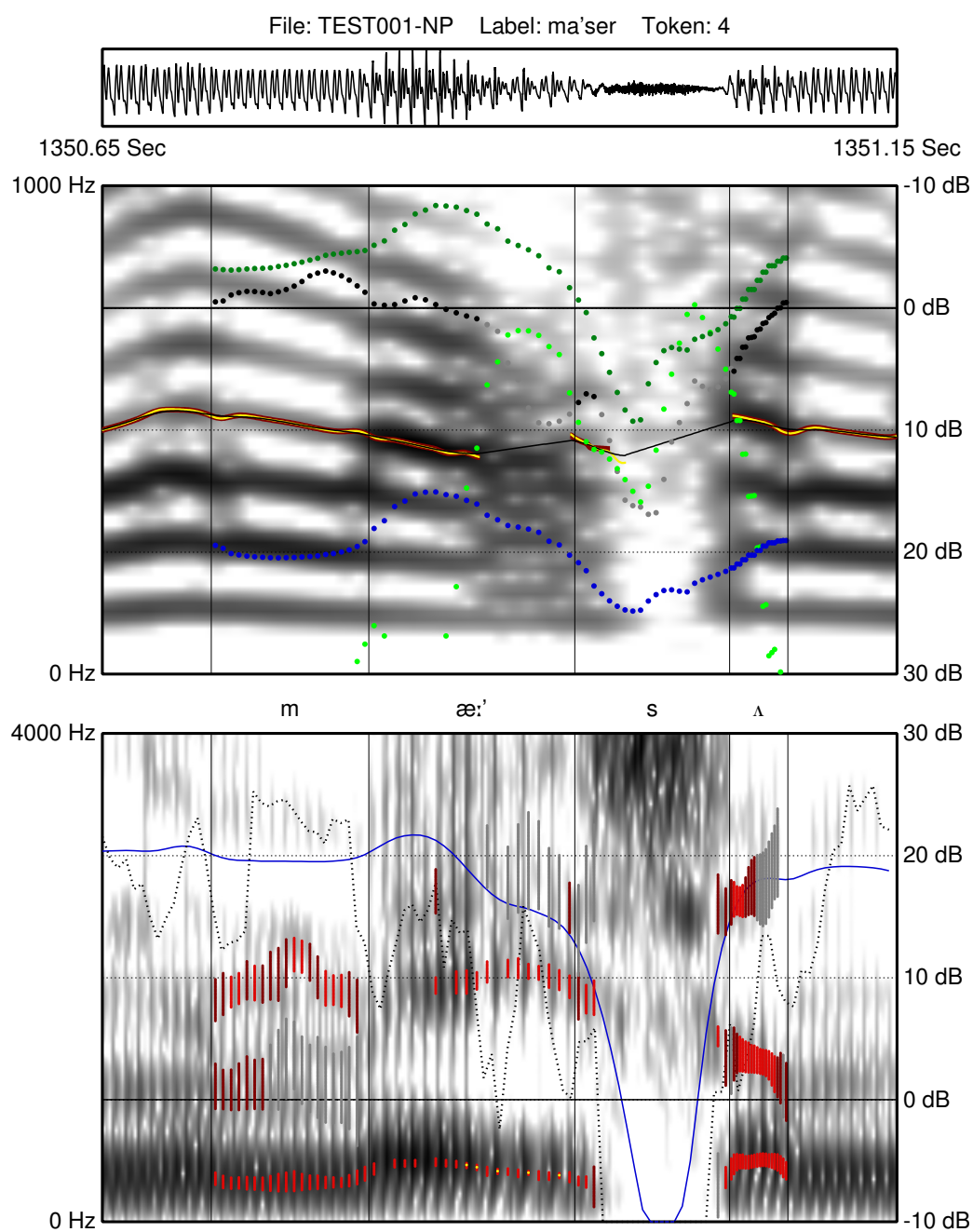
File: TEST001-NP Label: ma'ser Token: 2



H.III.2. ma'ser token 2

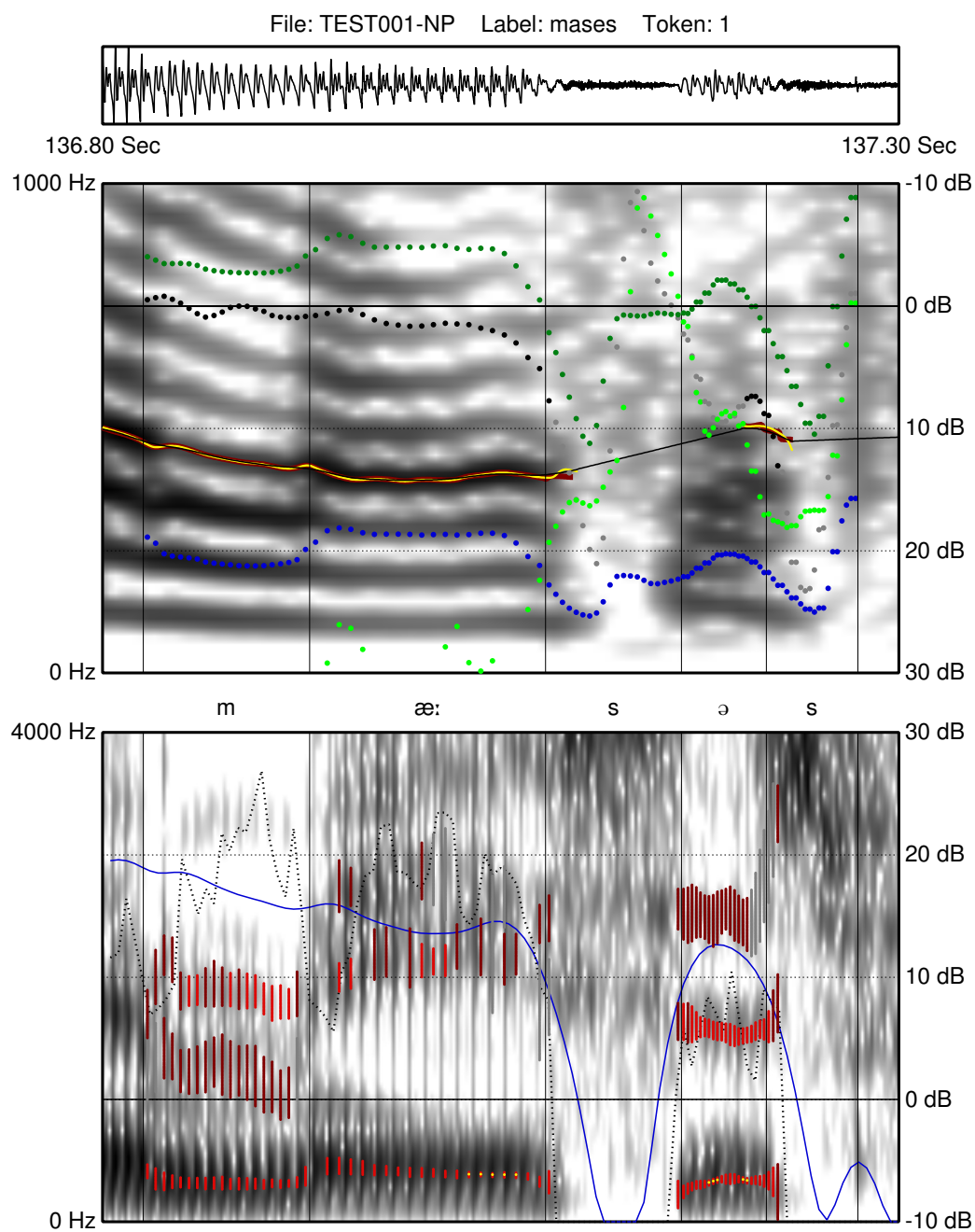


H.III.3. ma'ser token 3

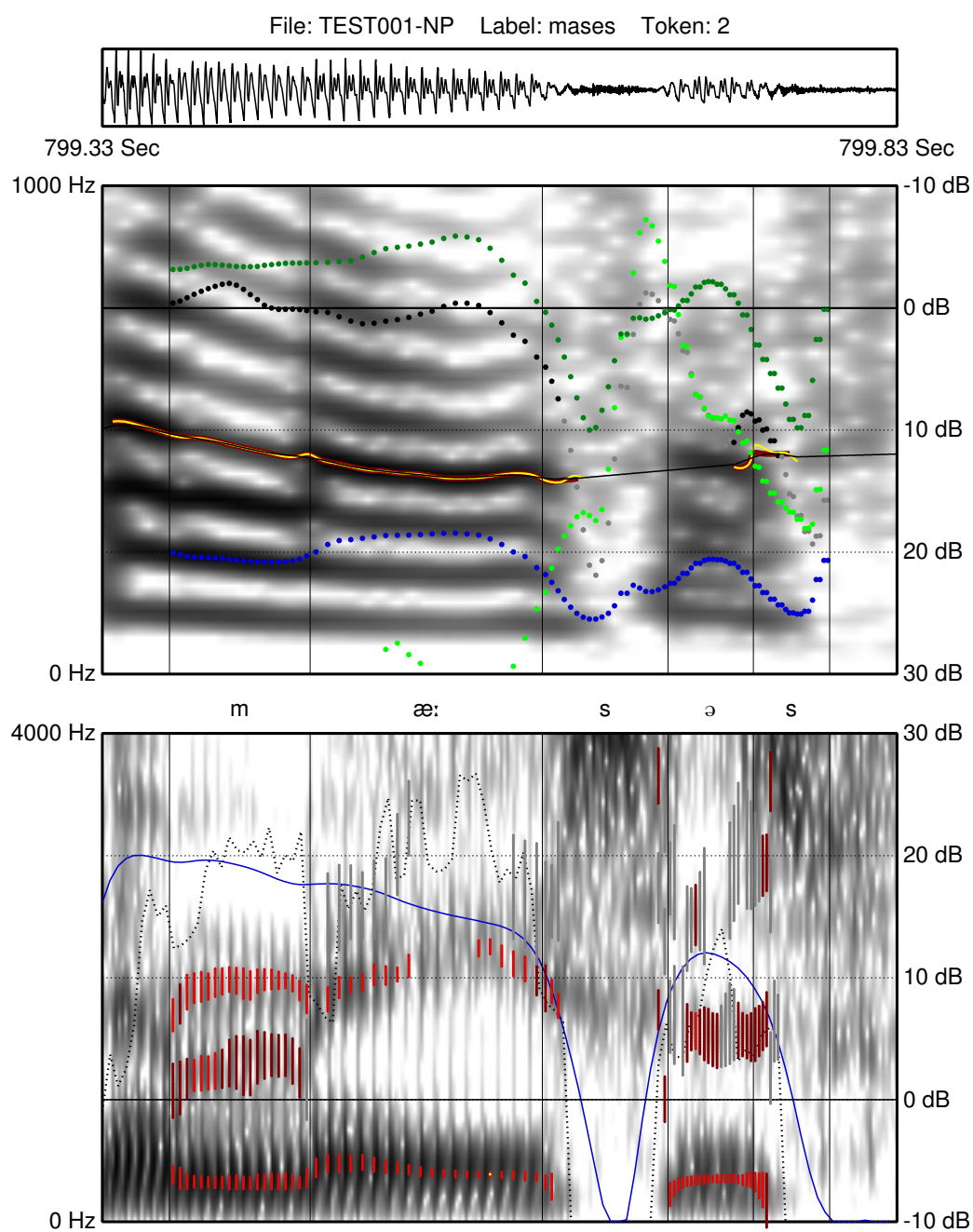


H.III.4. ma'ser token 4

H.IV mases

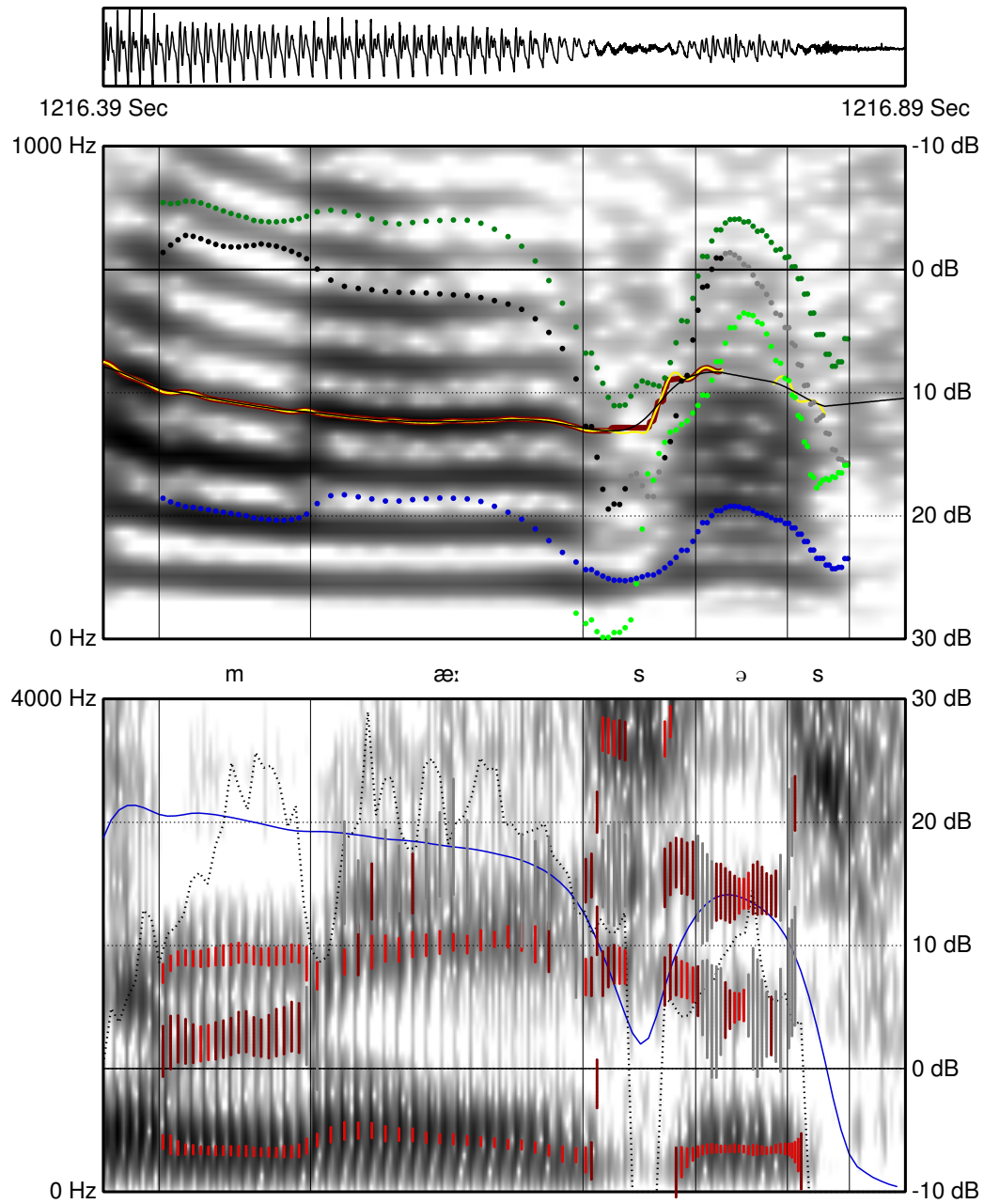


H.IV.1. mases token 1

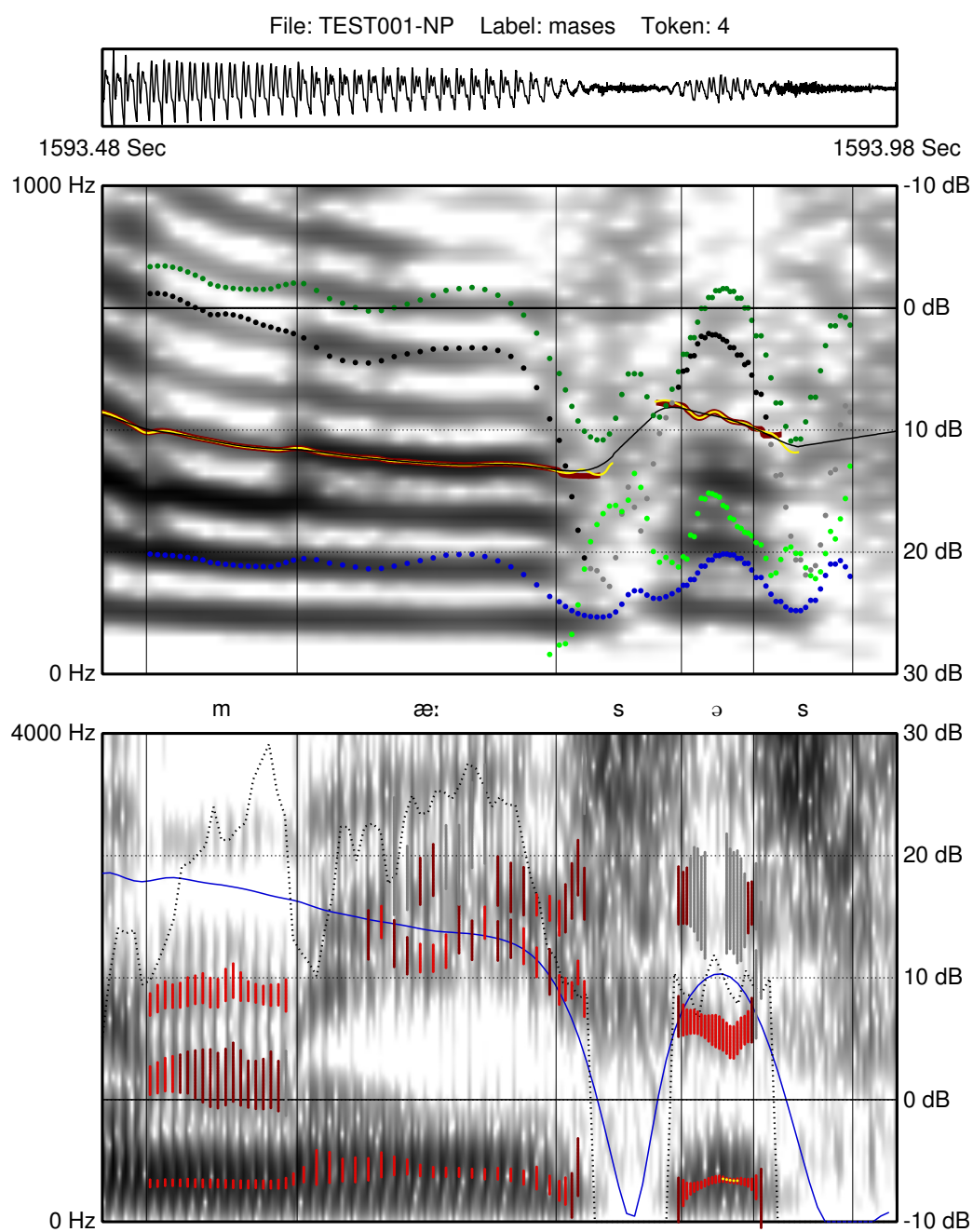


H.IV.2. mases token 2

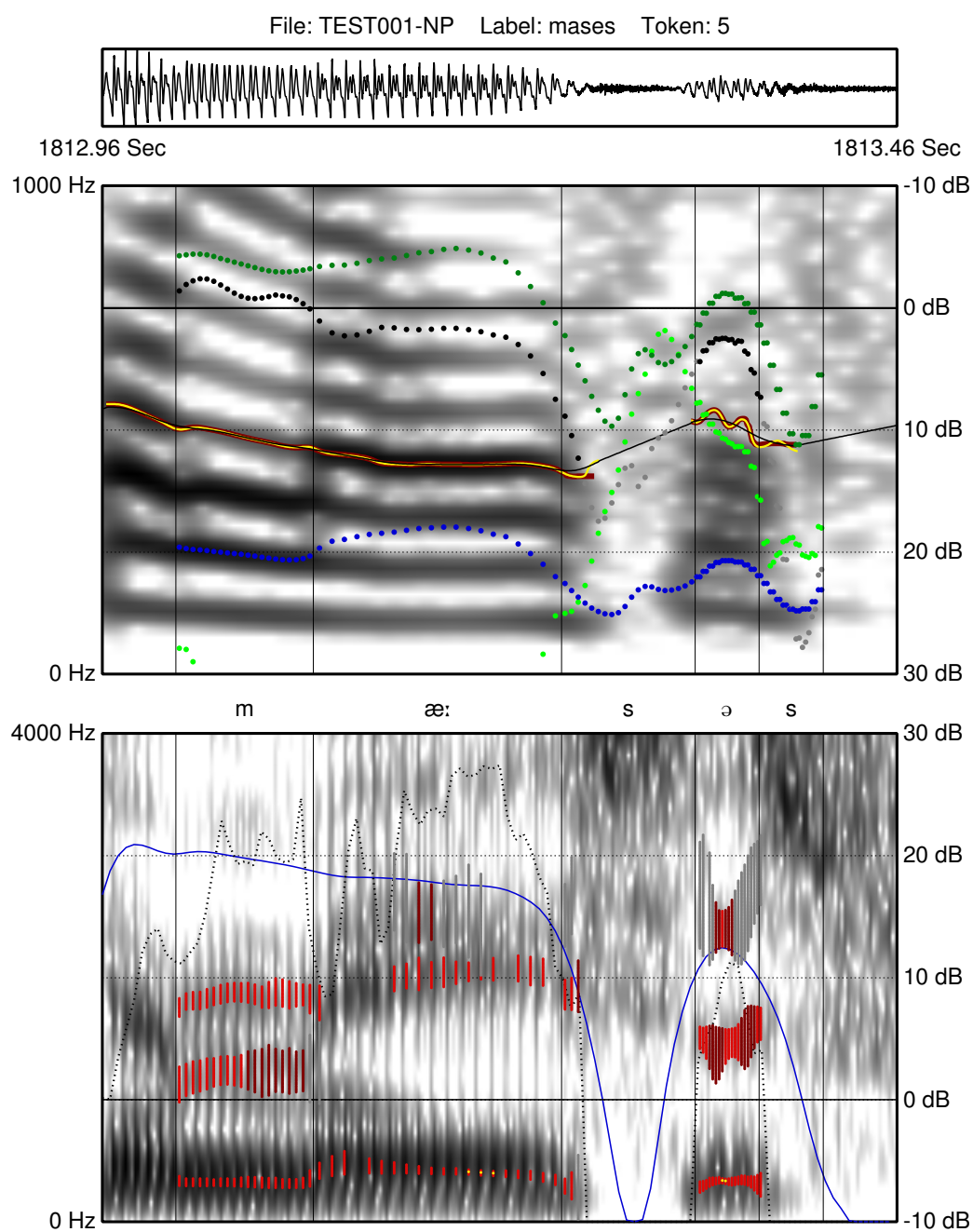
File: TEST001-NP Label: mases Token: 3



H.IV.3. mases token 3

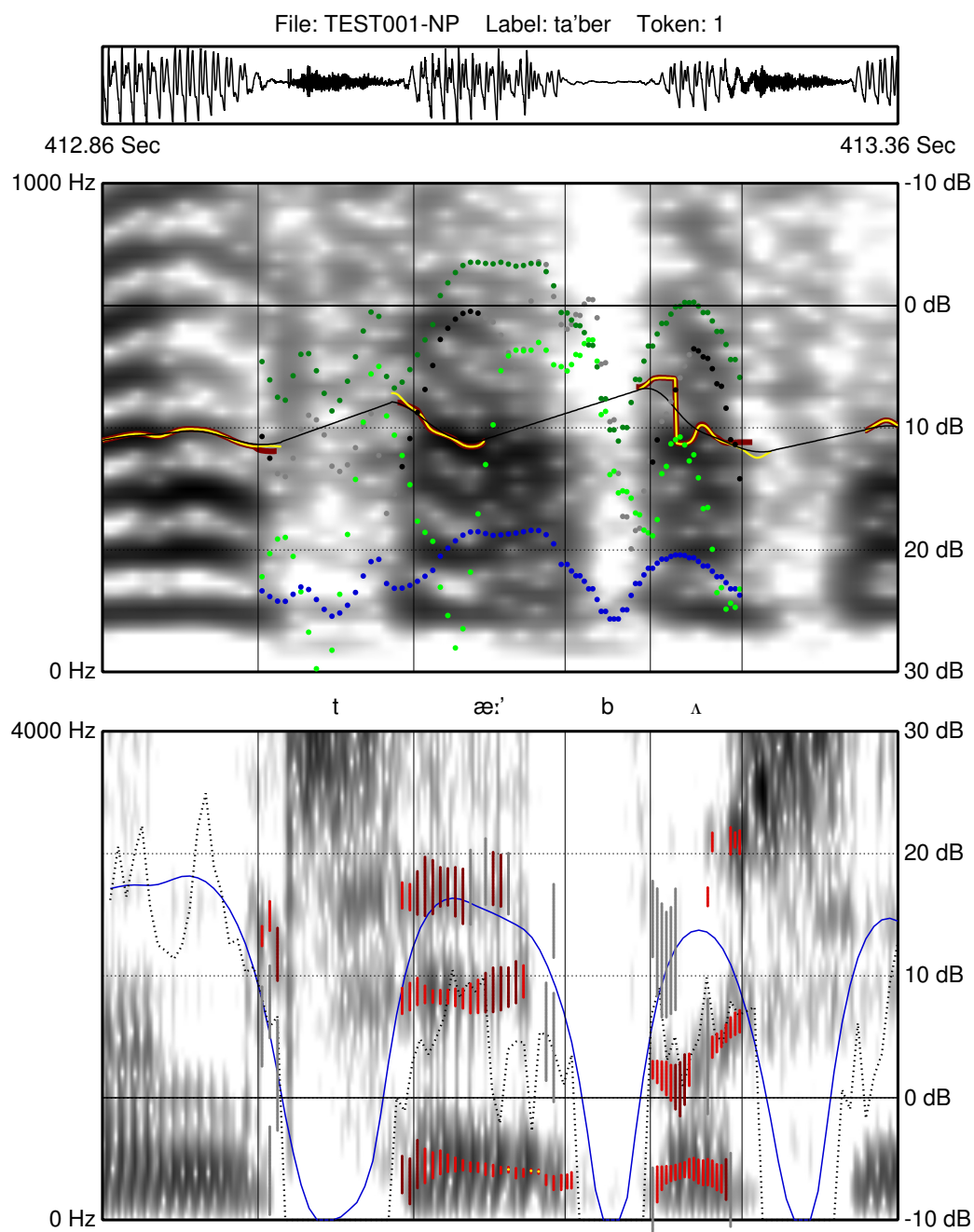


H.IV.4. mases token 4



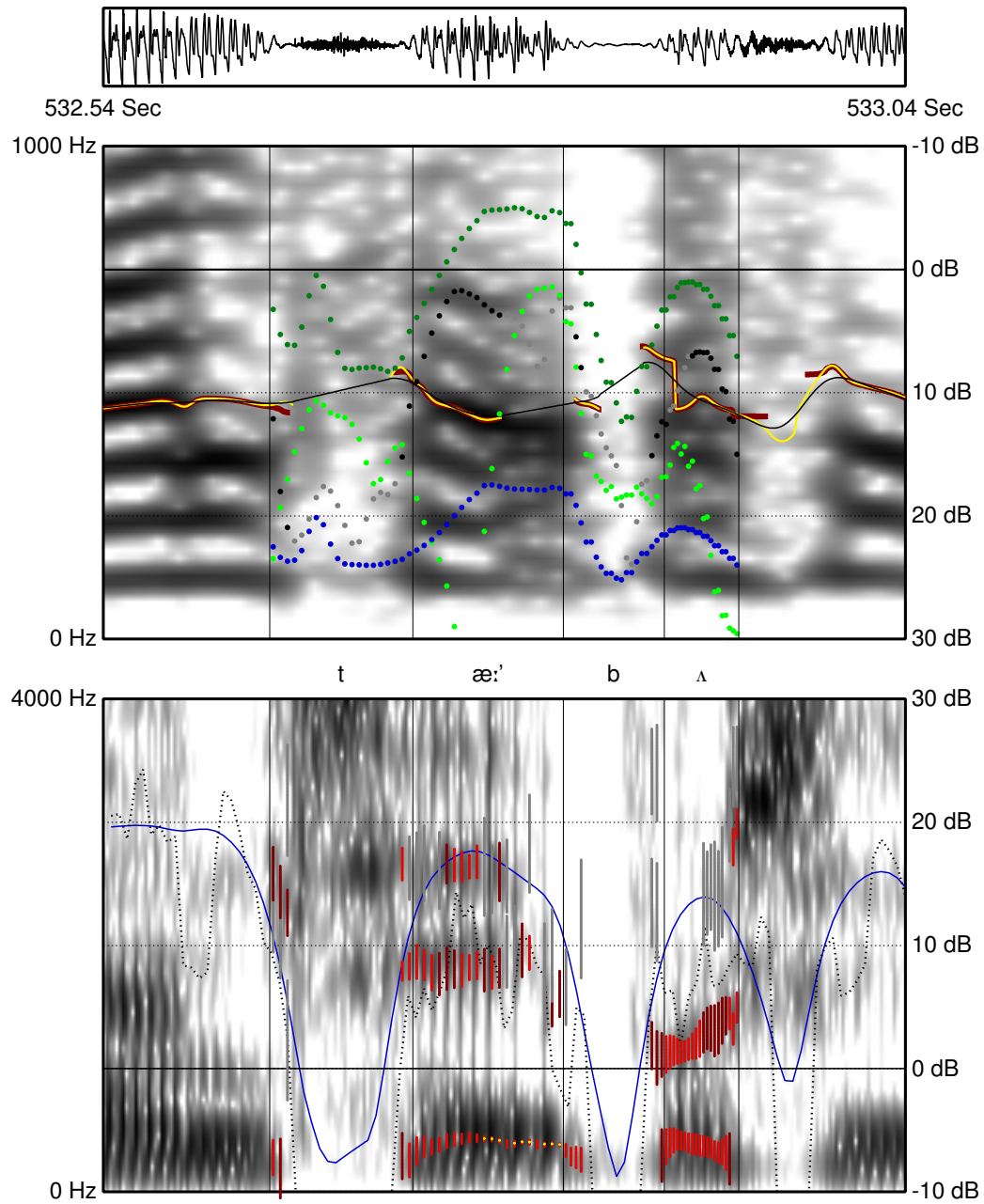
H.IV.5. mases token 5

H.V ta'ber

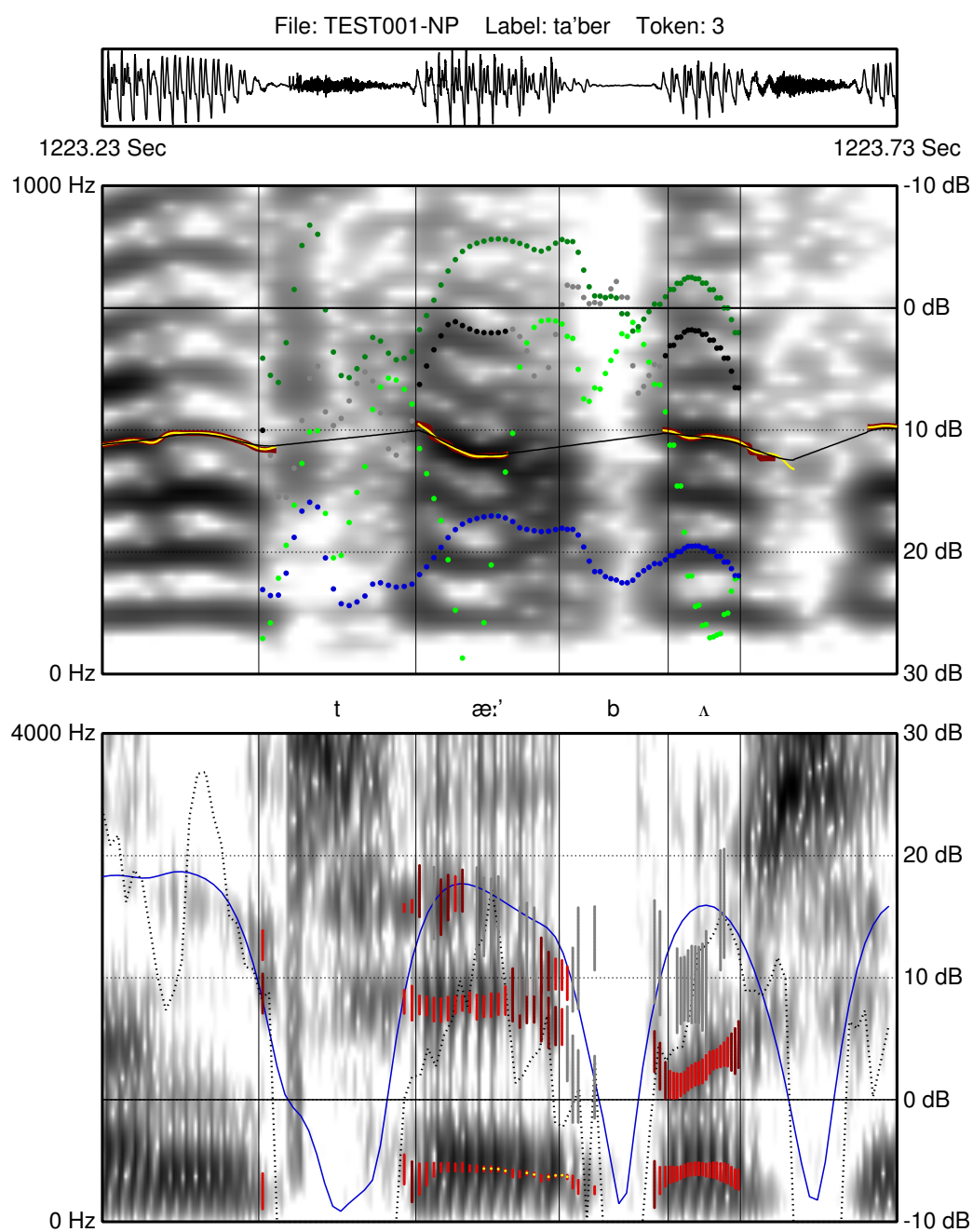


H.V.1. ta'ber token 1

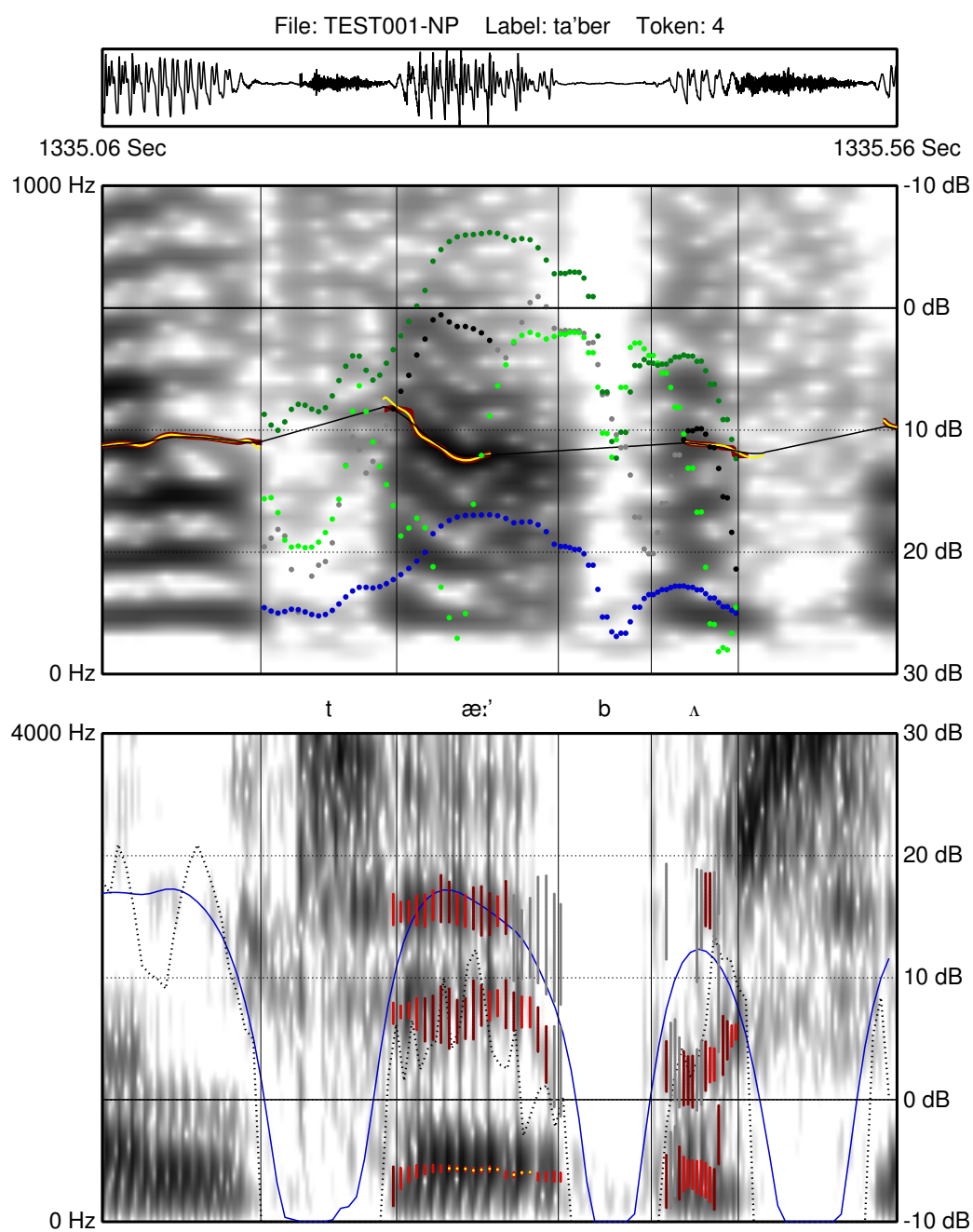
File: TEST001-NP Label: ta'ber Token: 2



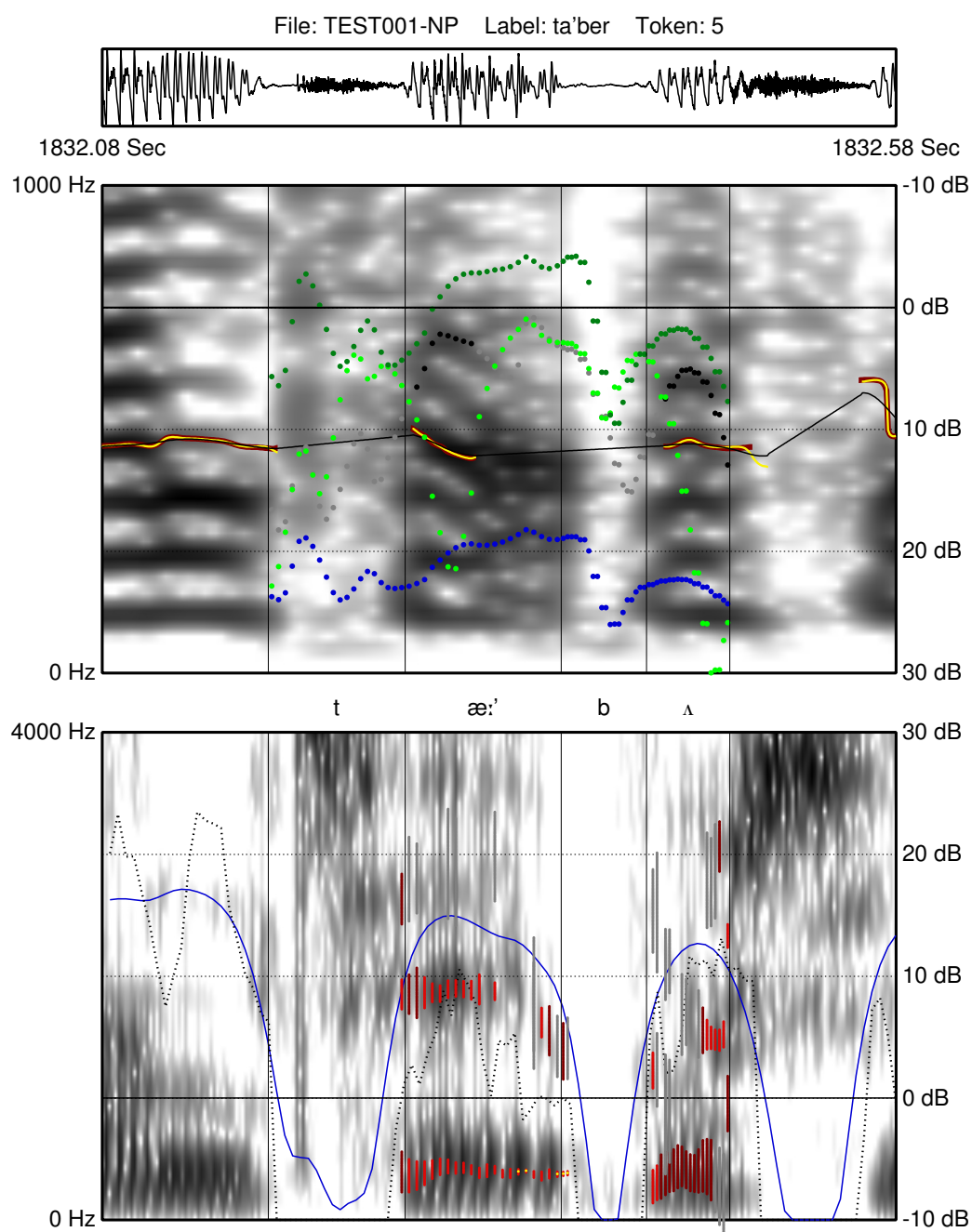
H.V.2. ta'ber token 2



H.V.3. ta'ber token 3

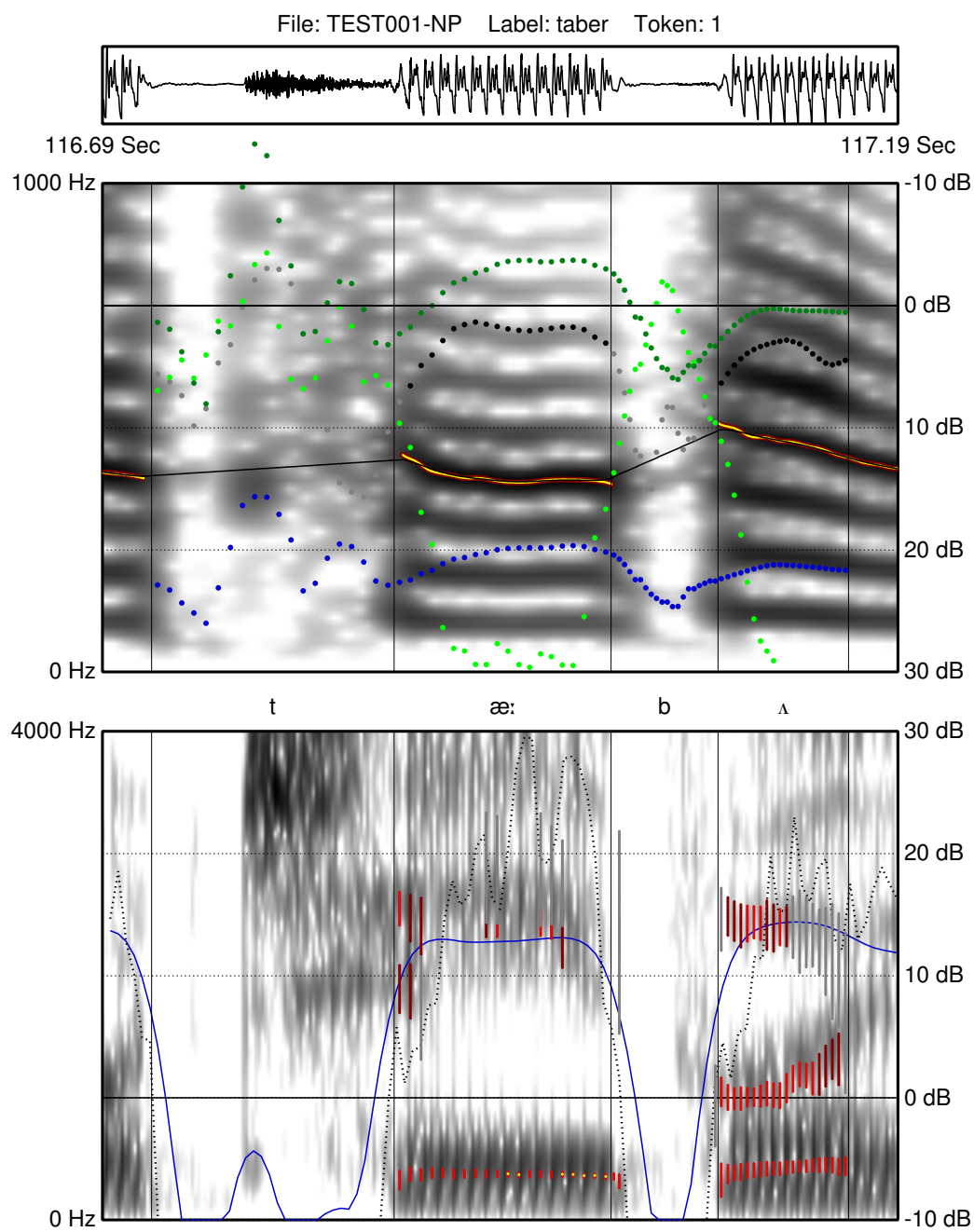


H.V.4. ta'ber token 4



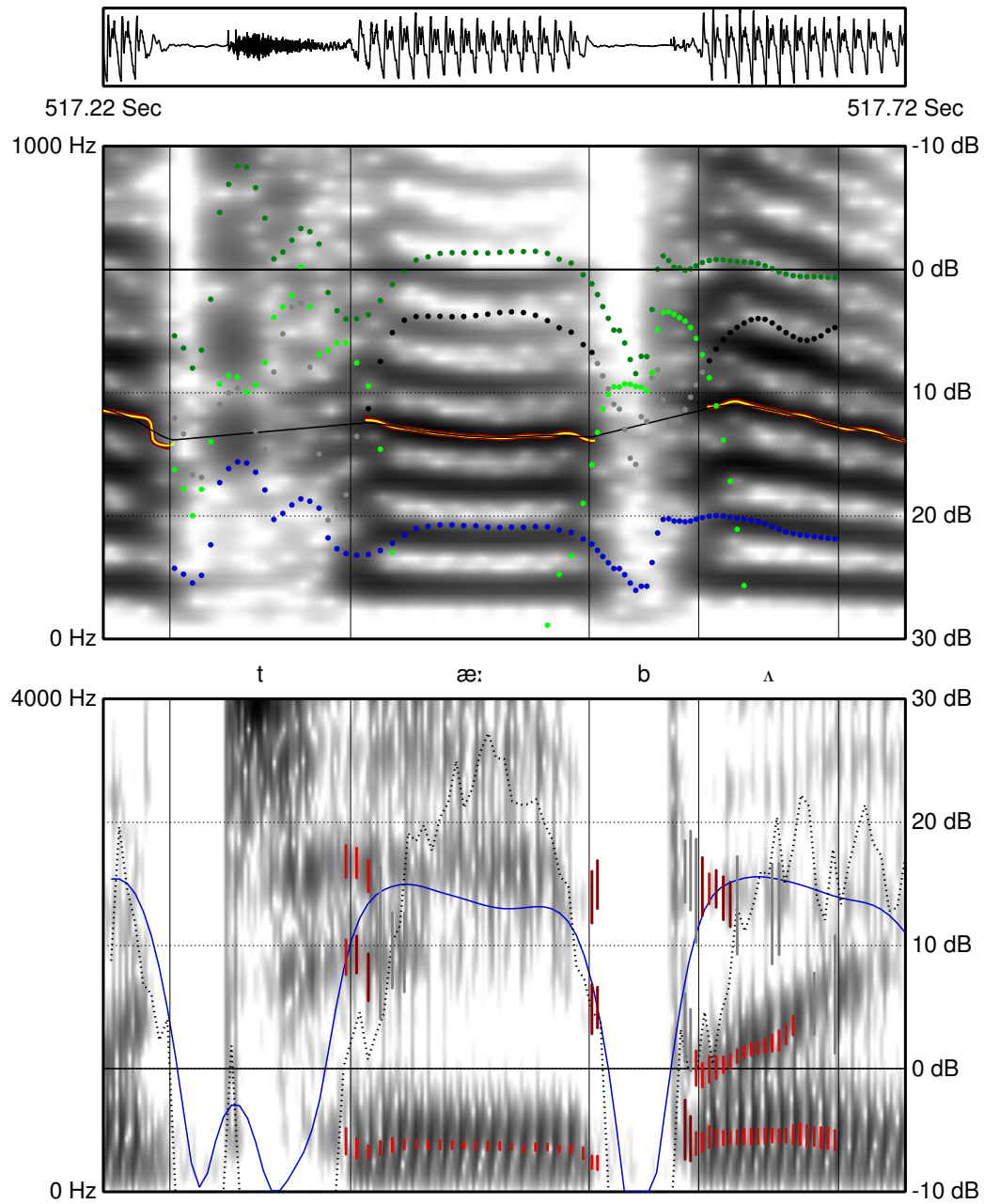
H.V.5. ta'ber token 5

H.VI taber



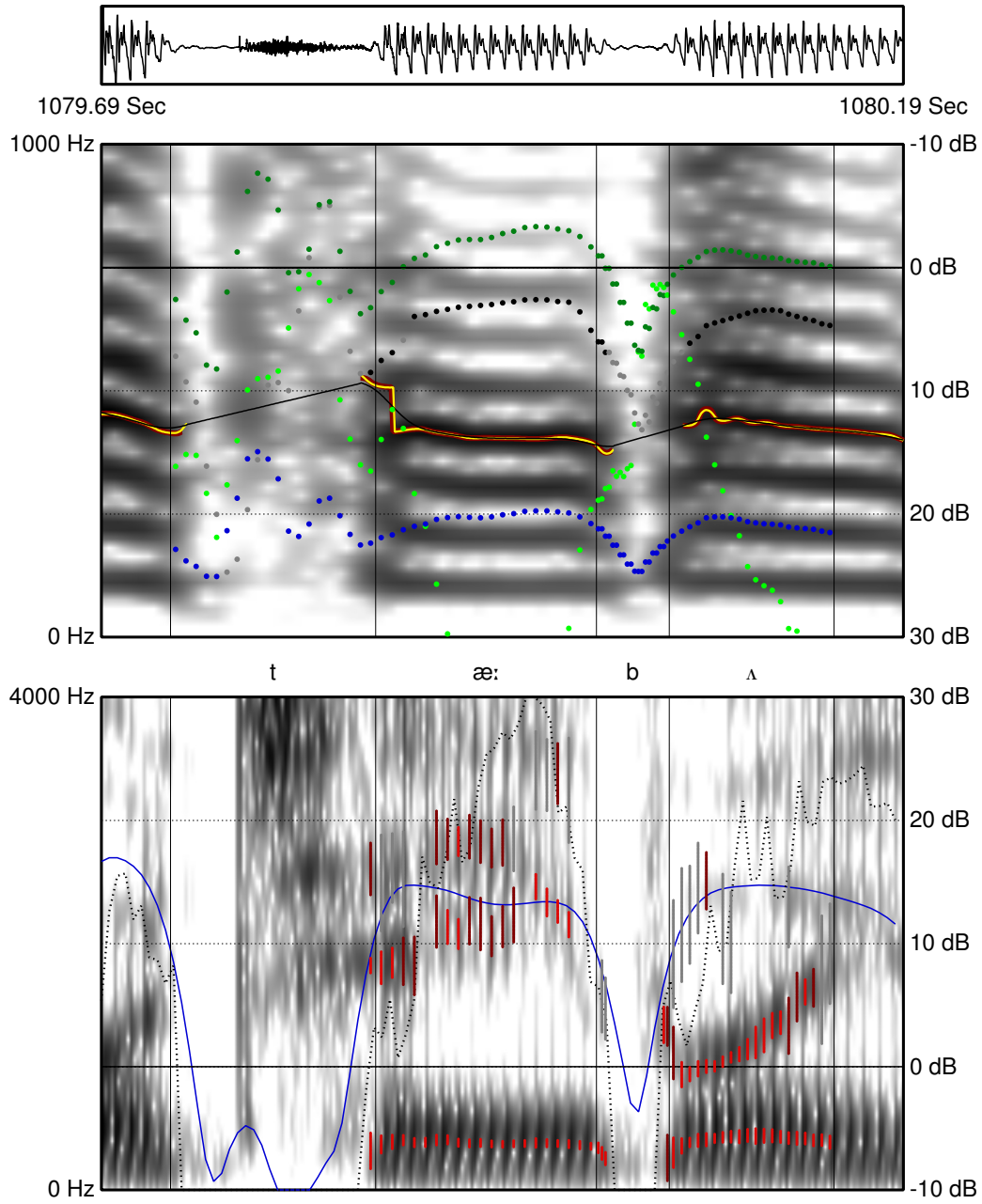
H.VI.1. taber token 1

File: TEST001-NP Label: taber Token: 2



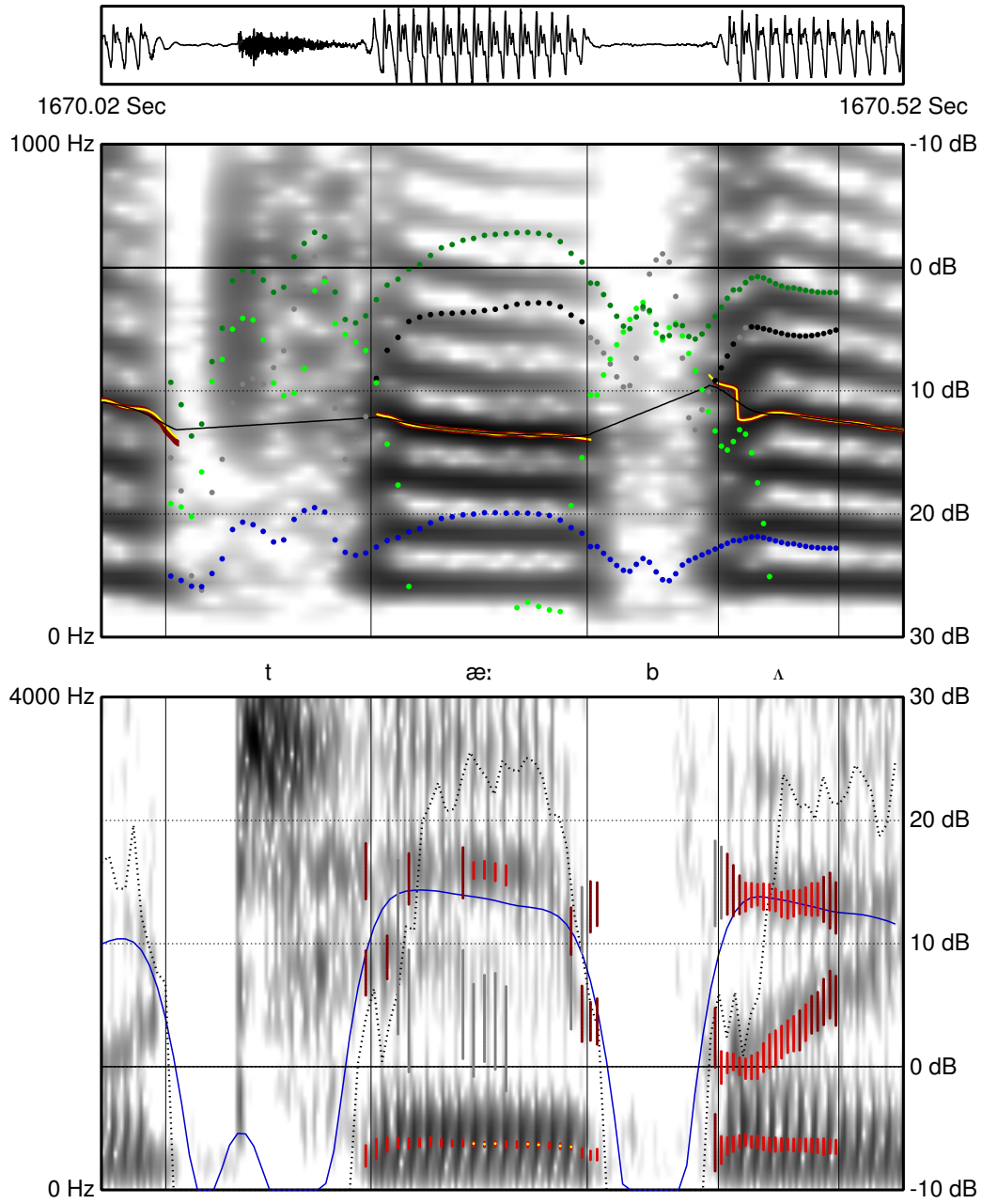
H.VI.2. taber token 2

File: TEST001-NP Label: taber Token: 3



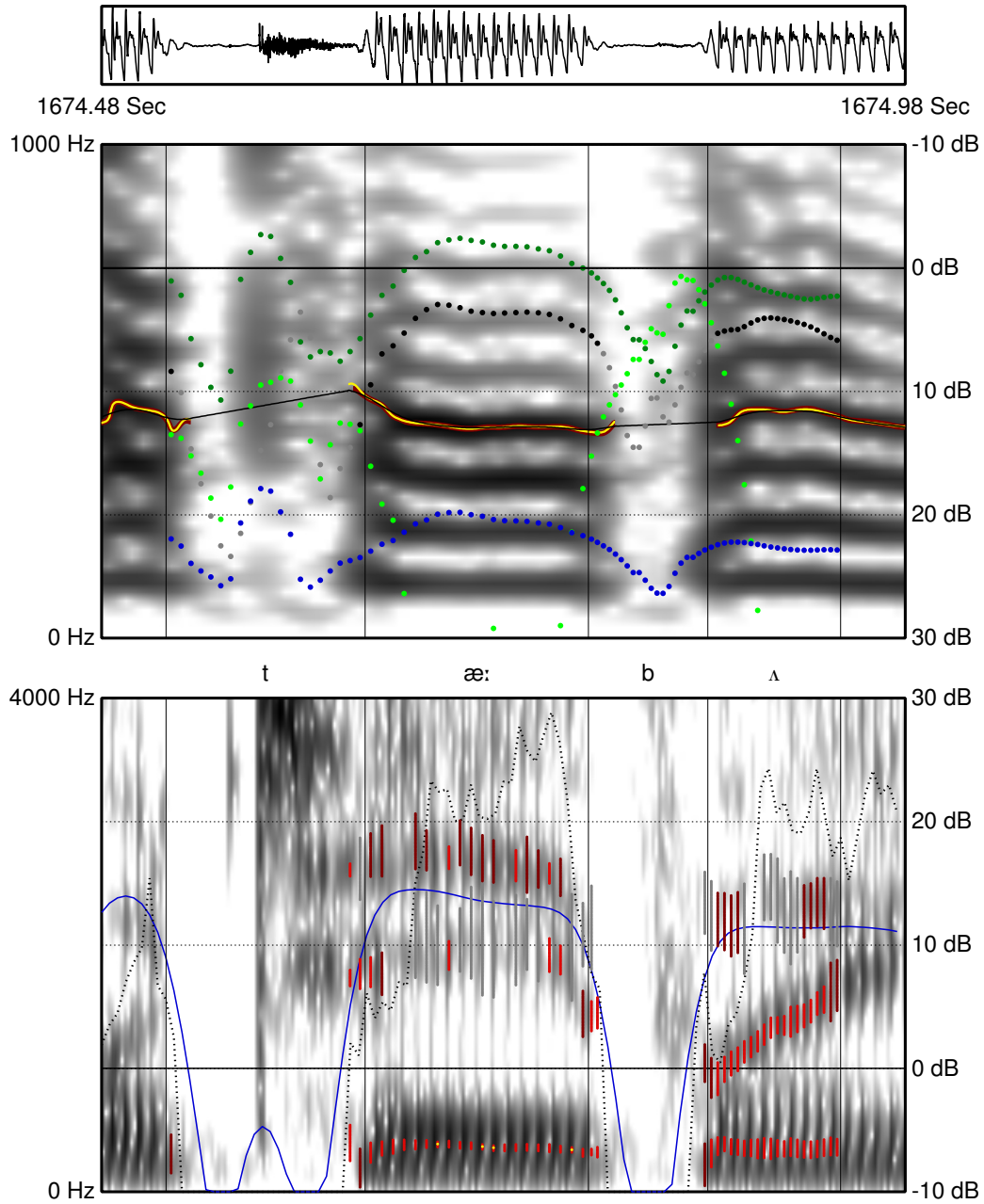
H.VI.3. taber token 3

File: TEST001-NP Label: taber Token: 4



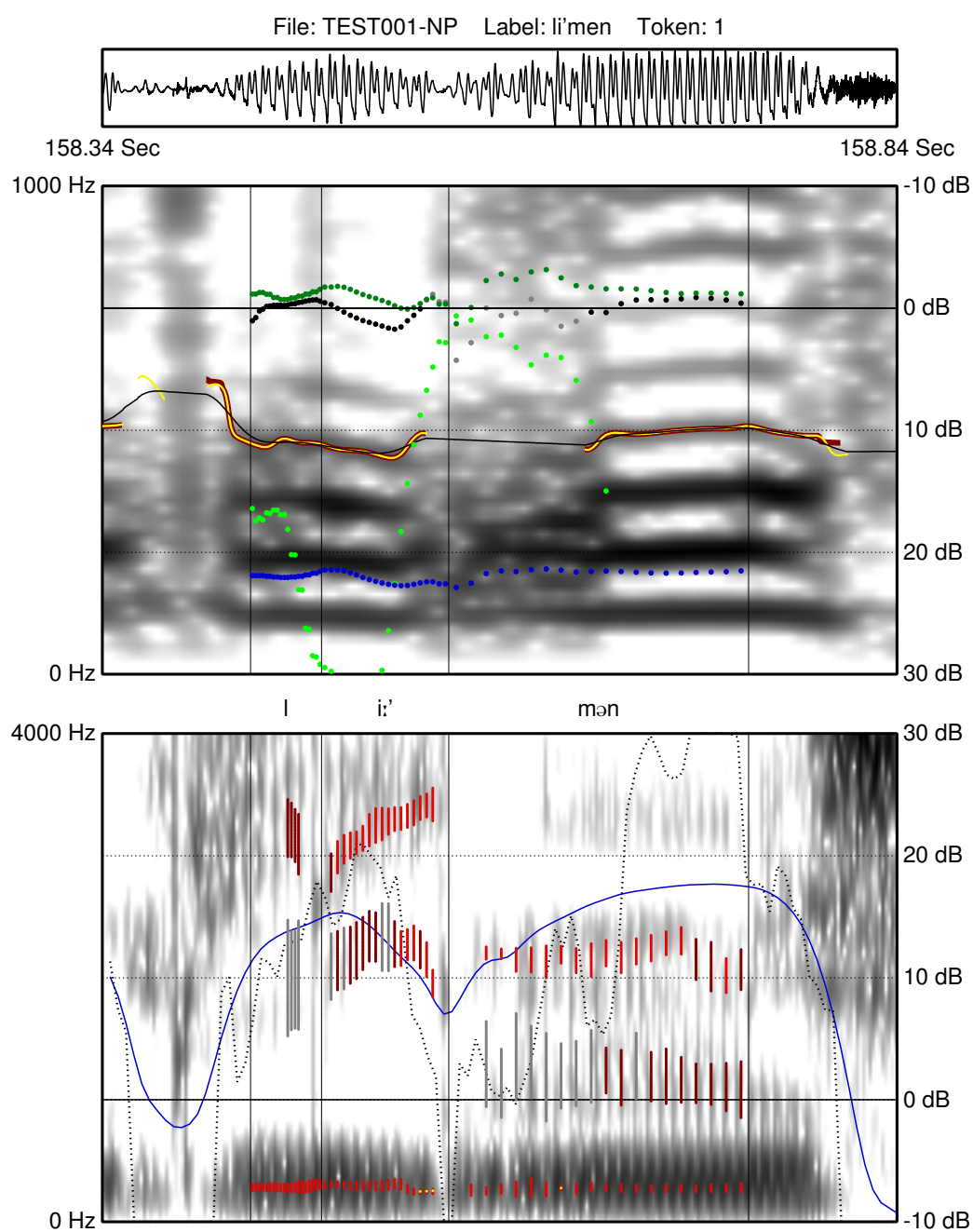
H.VI.4. taber token 4

File: TEST001-NP Label: taber Token: 5

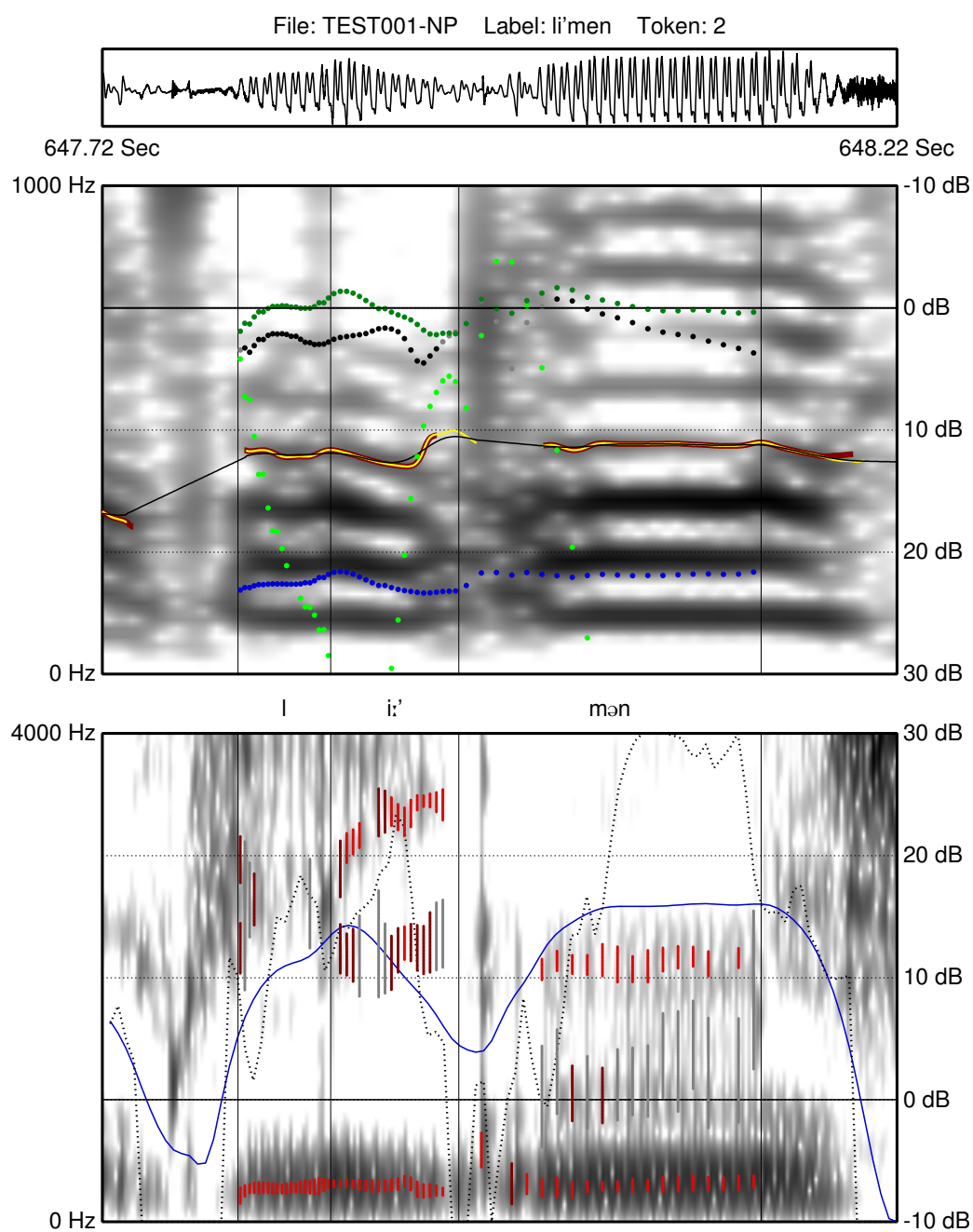


H.VI.5. taber token 5

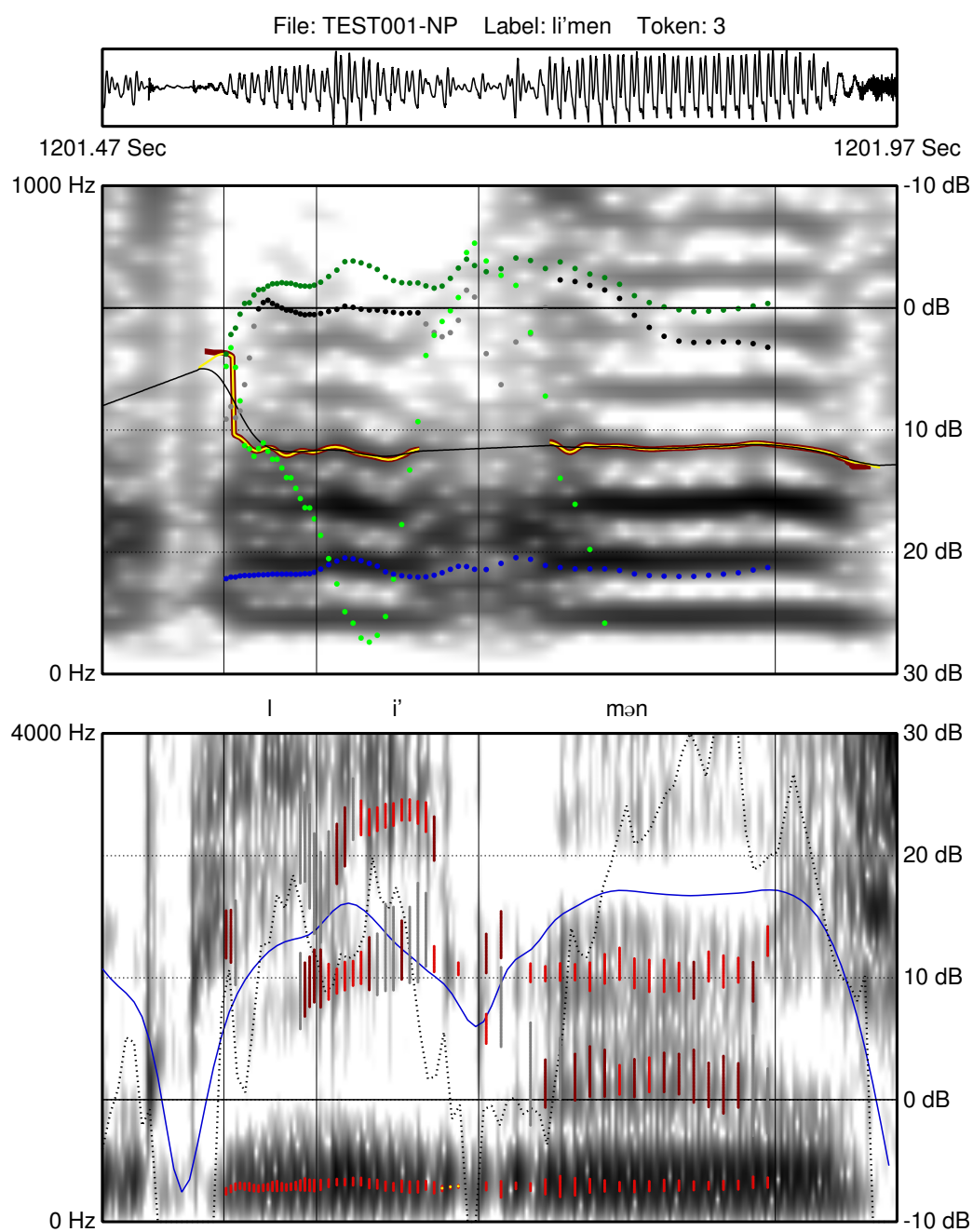
H.VII li'men



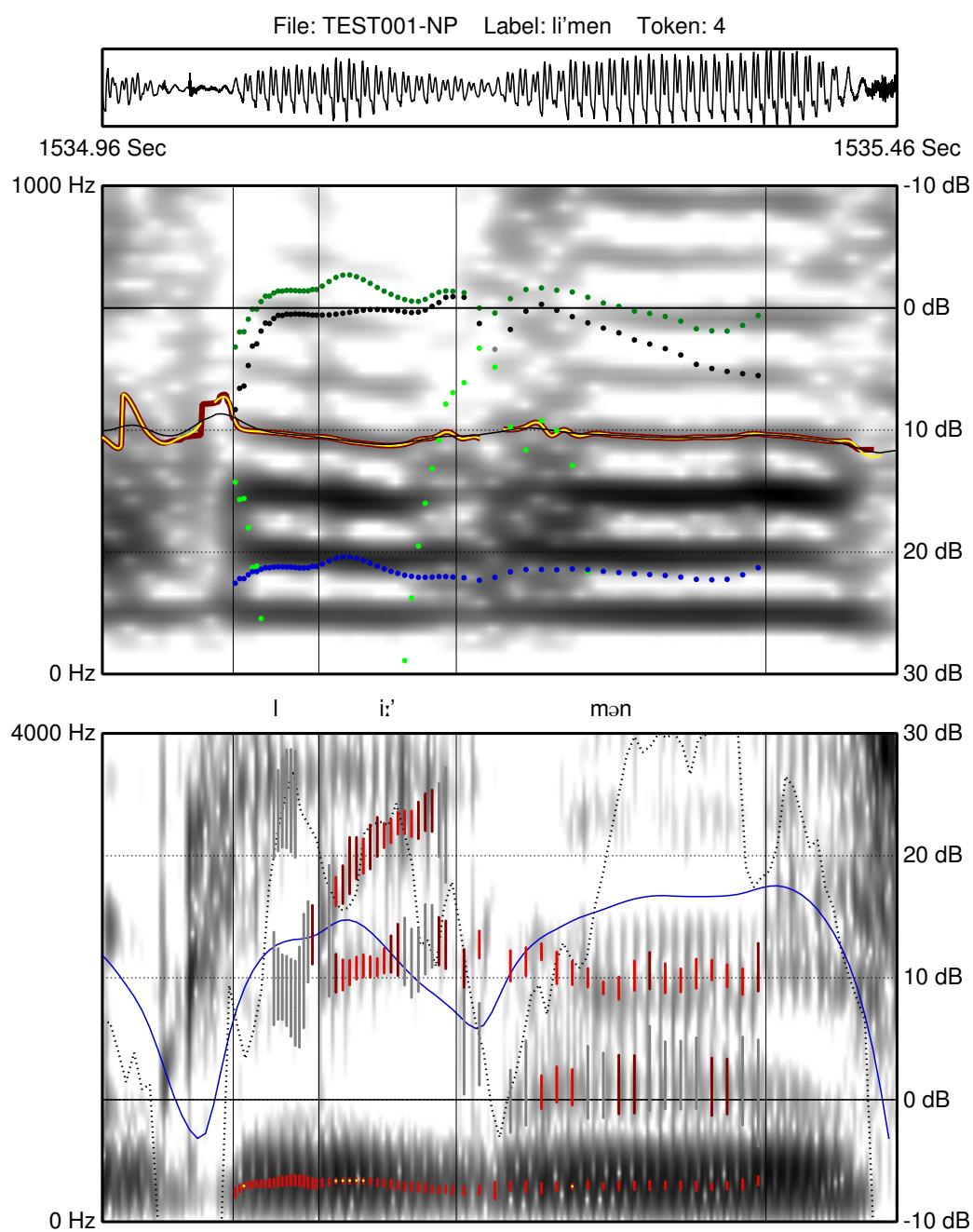
H.VII.1. li'men token 1



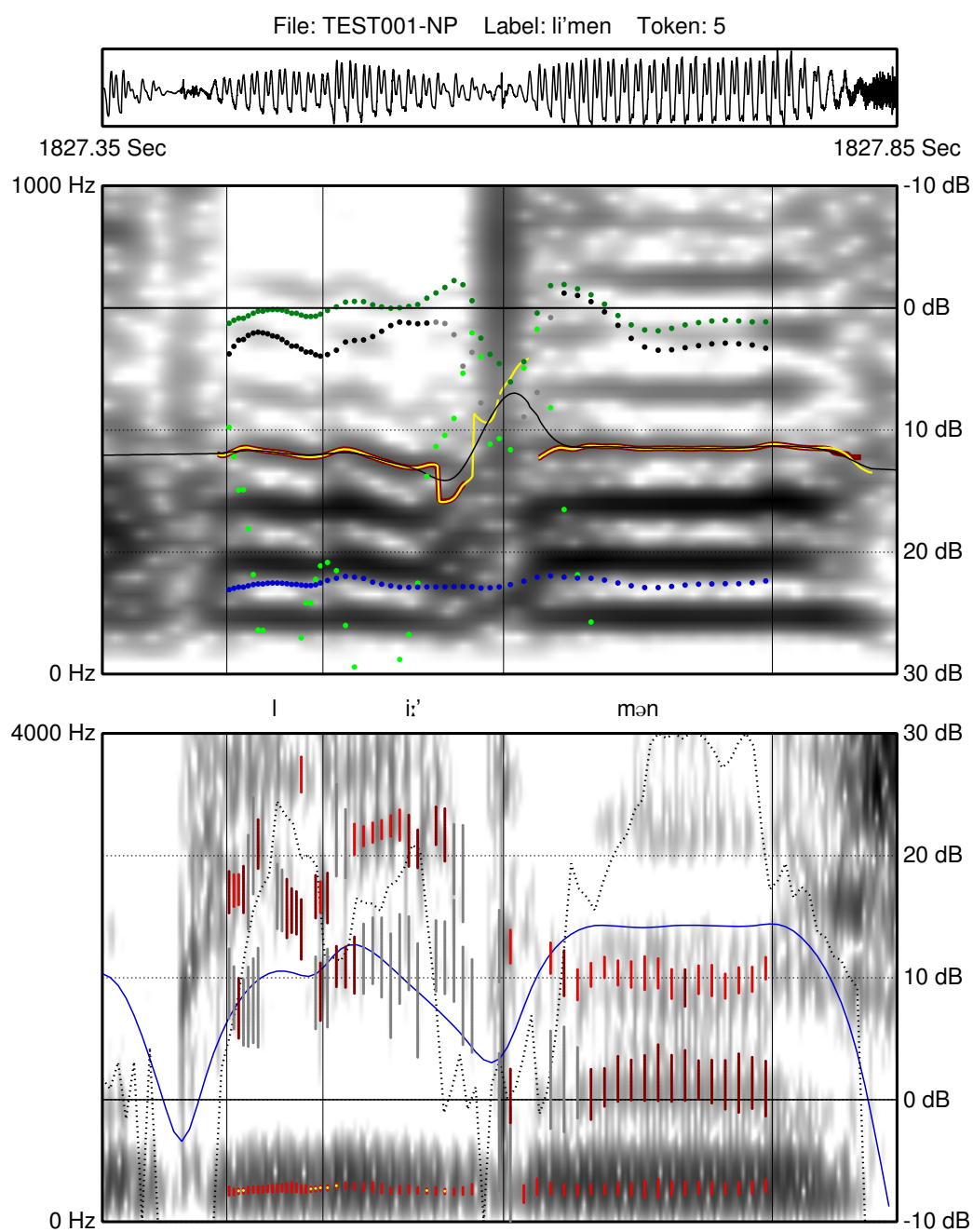
H.VII.2. li'men token 2



H.VII.3. li'men token 3

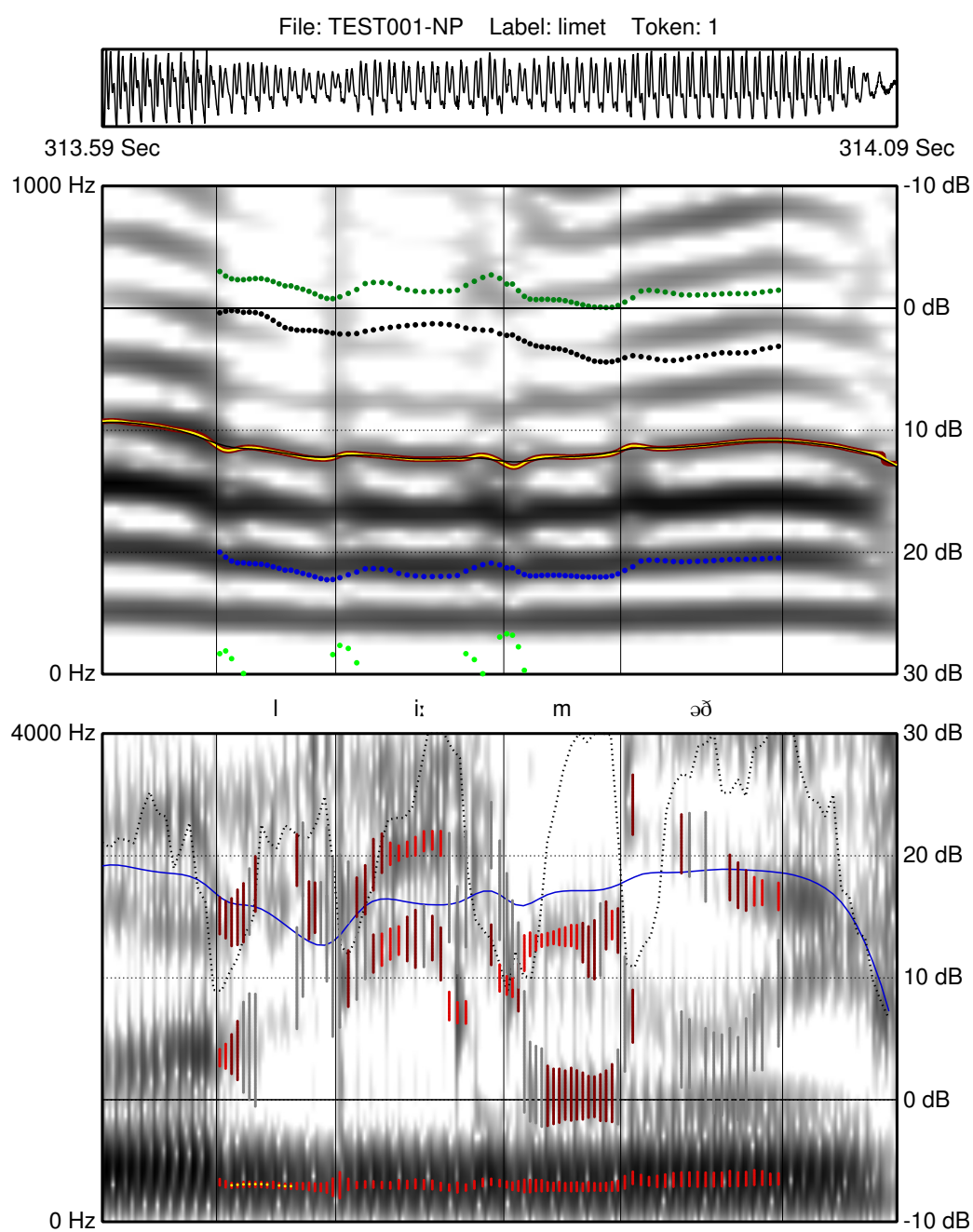


H.VII.4. li'men token 4

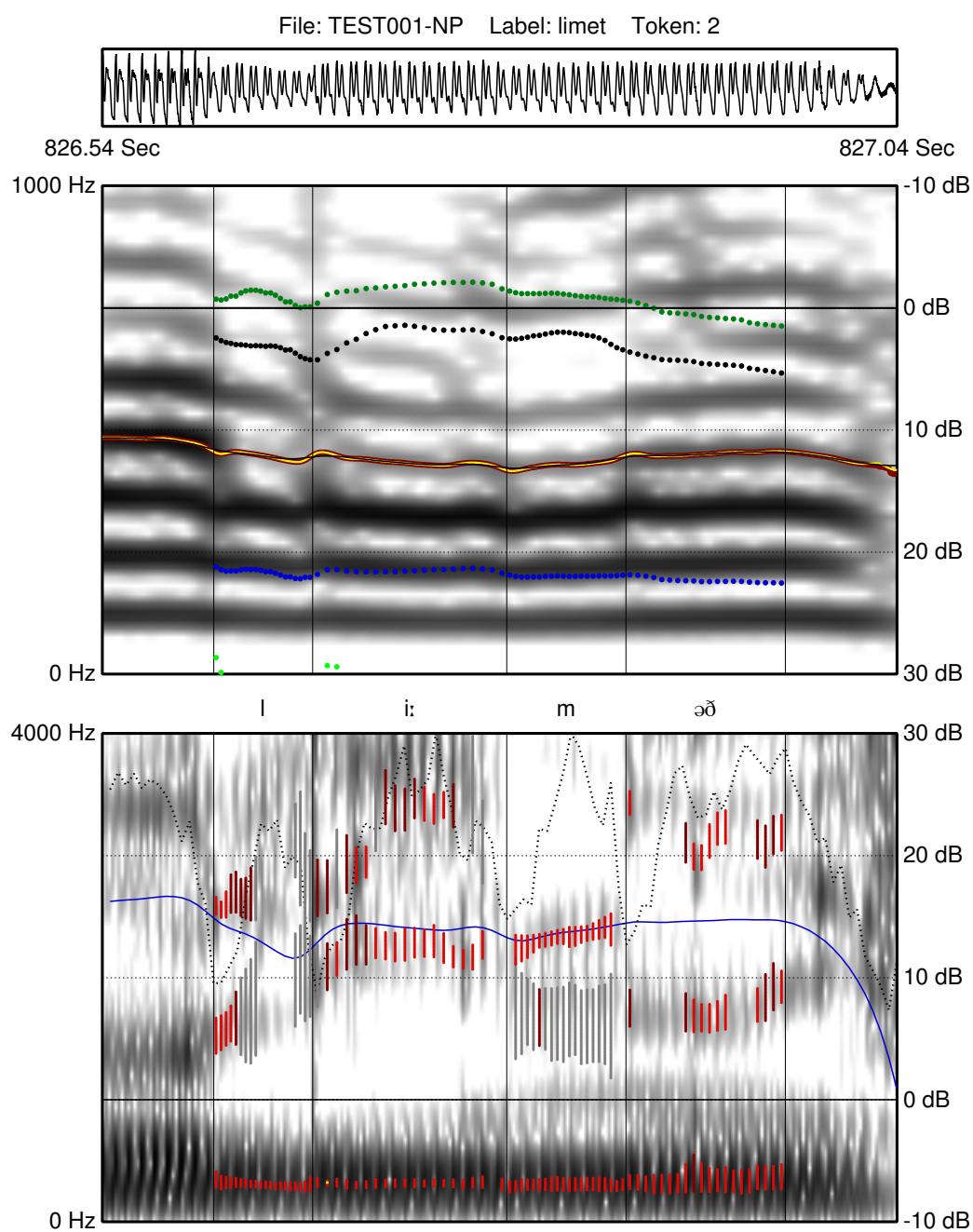


H.VII.5. li'men token 5

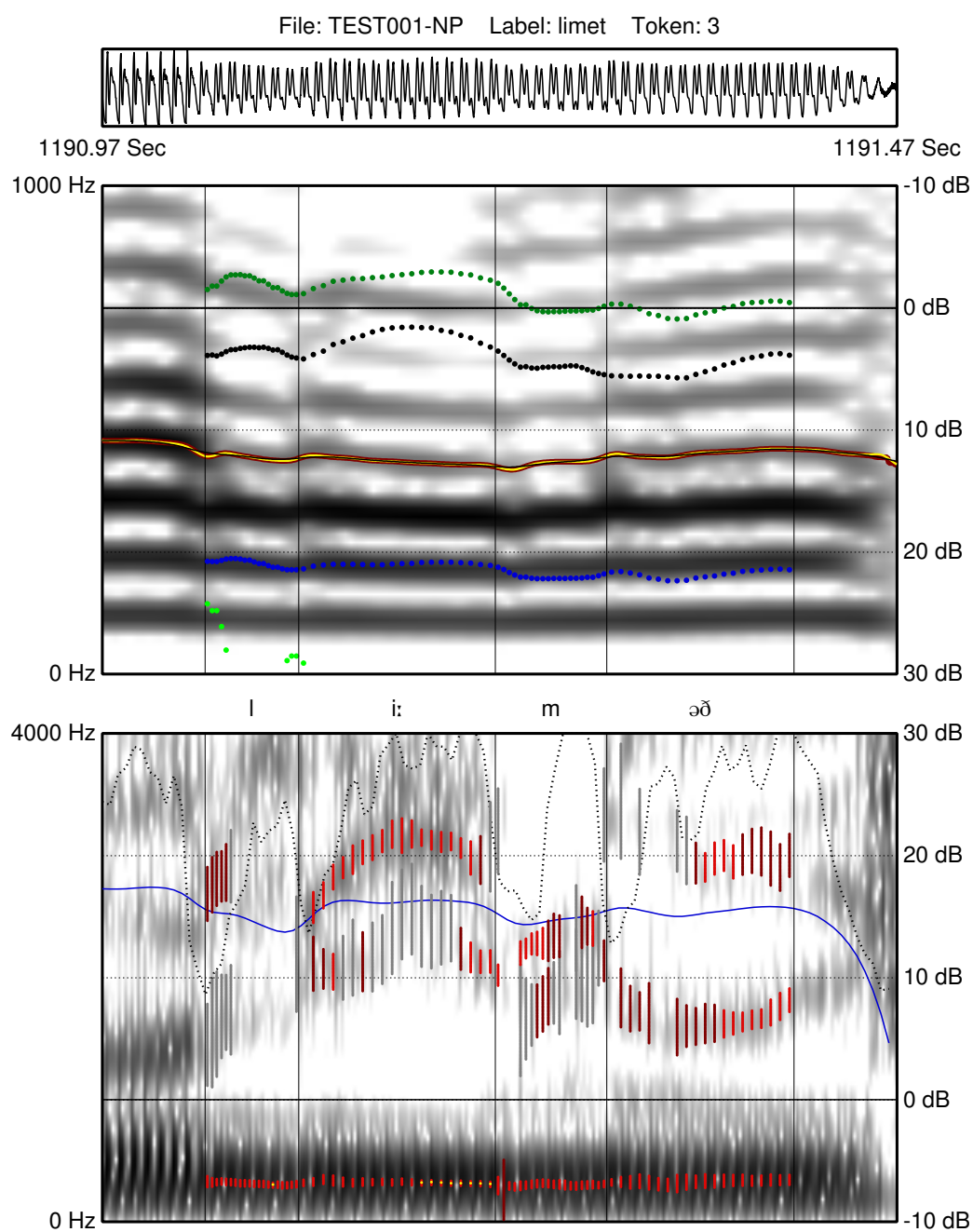
H.VIII limet



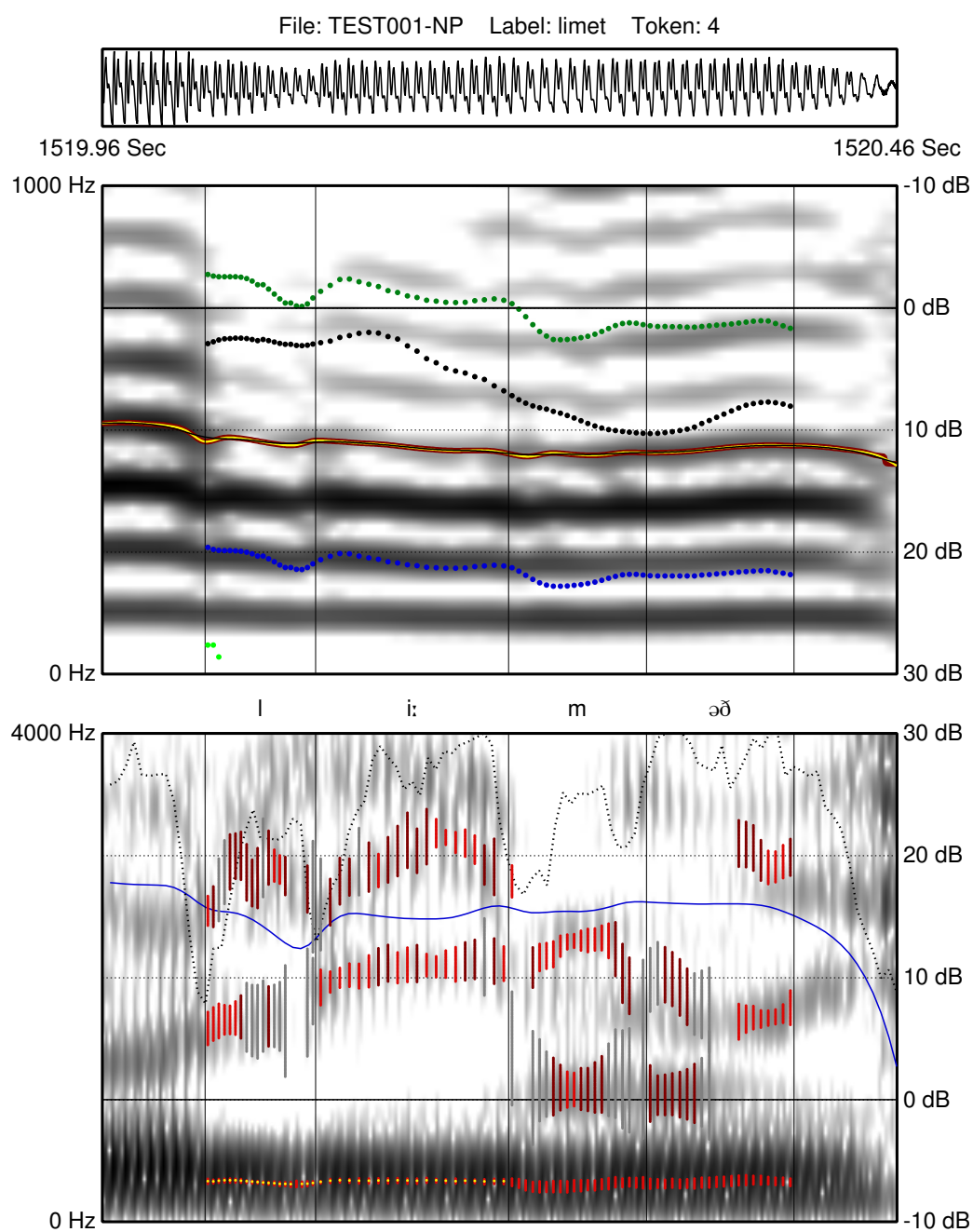
H.VIII.1. limet token 1



H.VIII.2. limet token 2

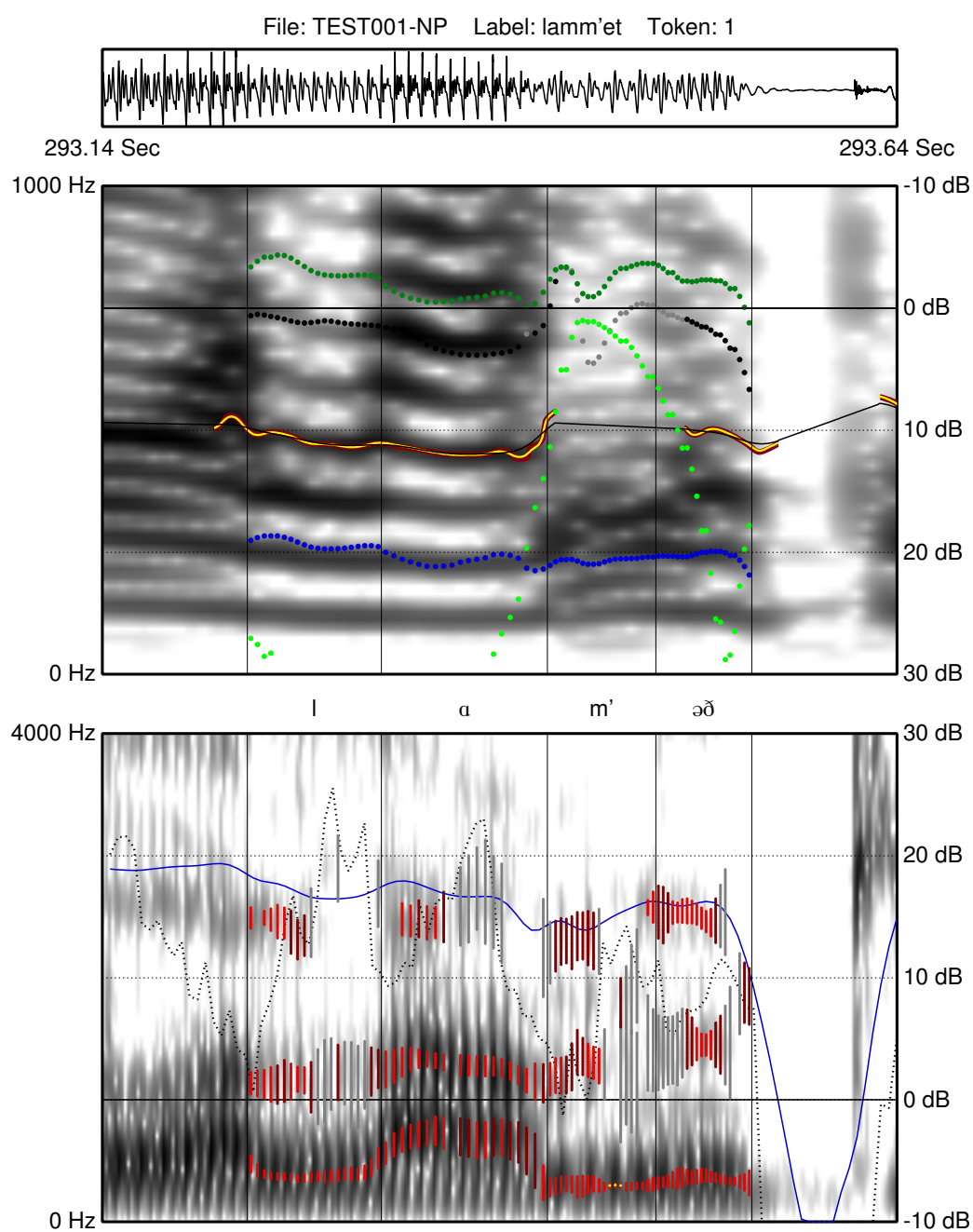


H.VIII.3. limet token 3

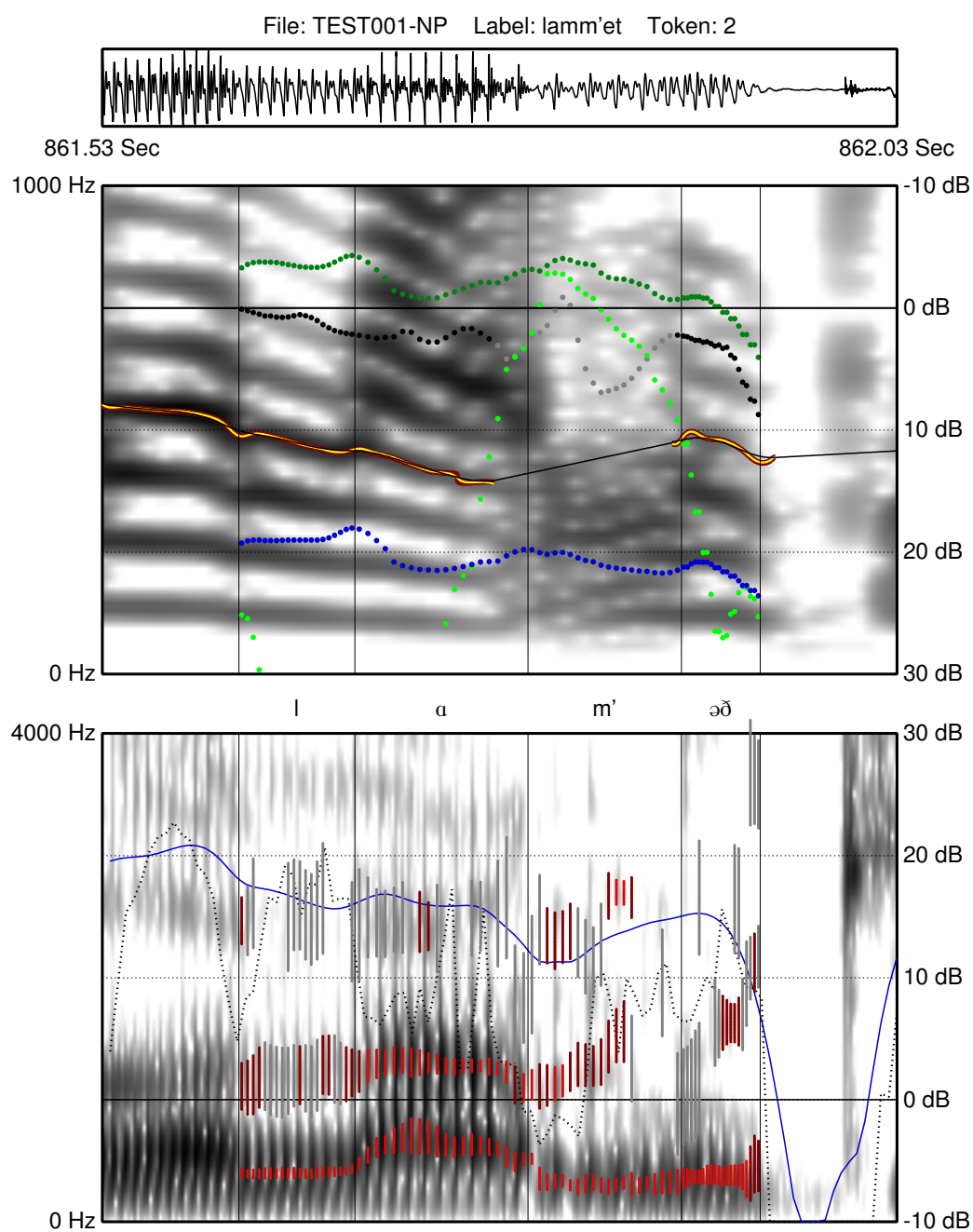


H.VIII.4. limet token 4

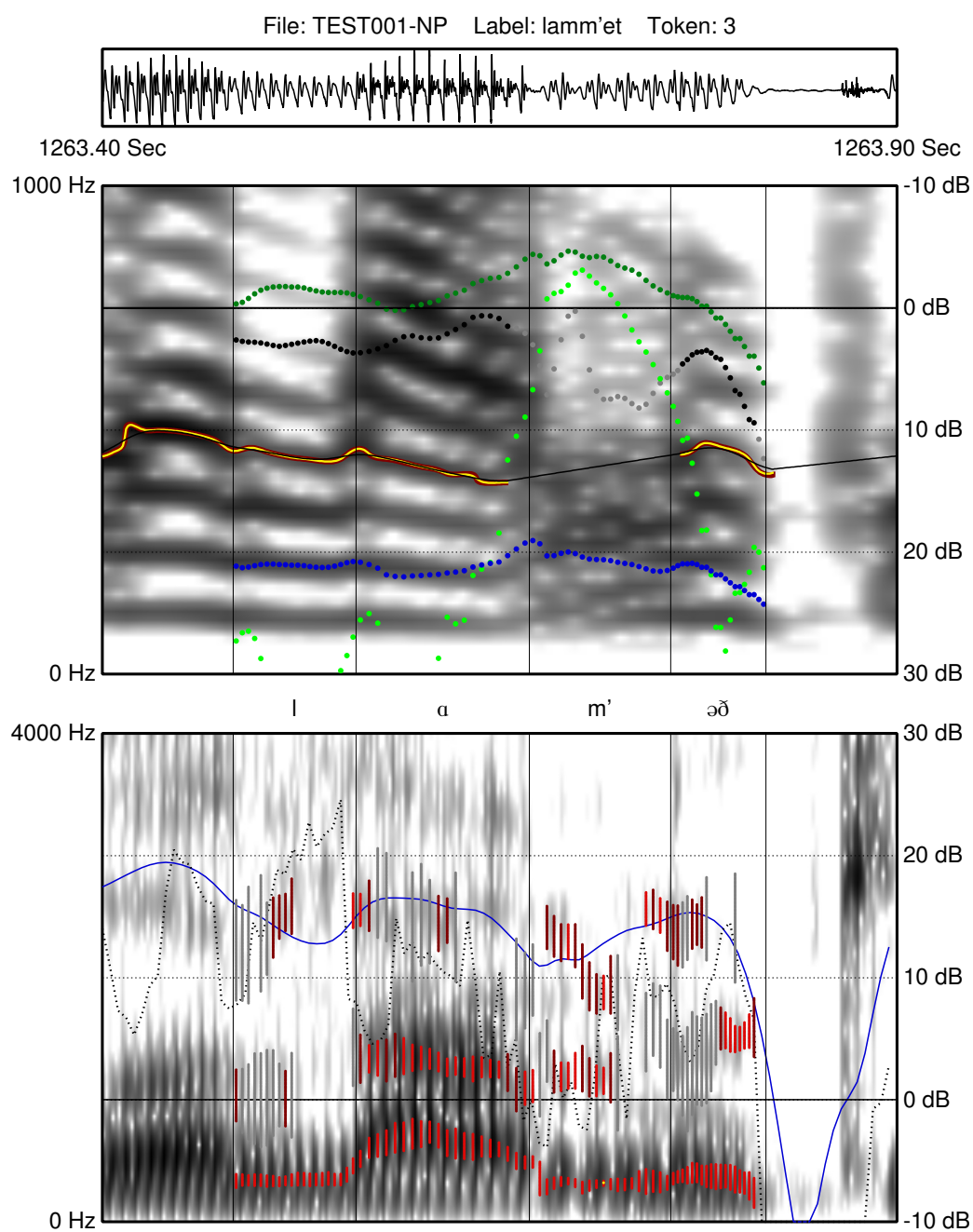
H.IX lamm'et



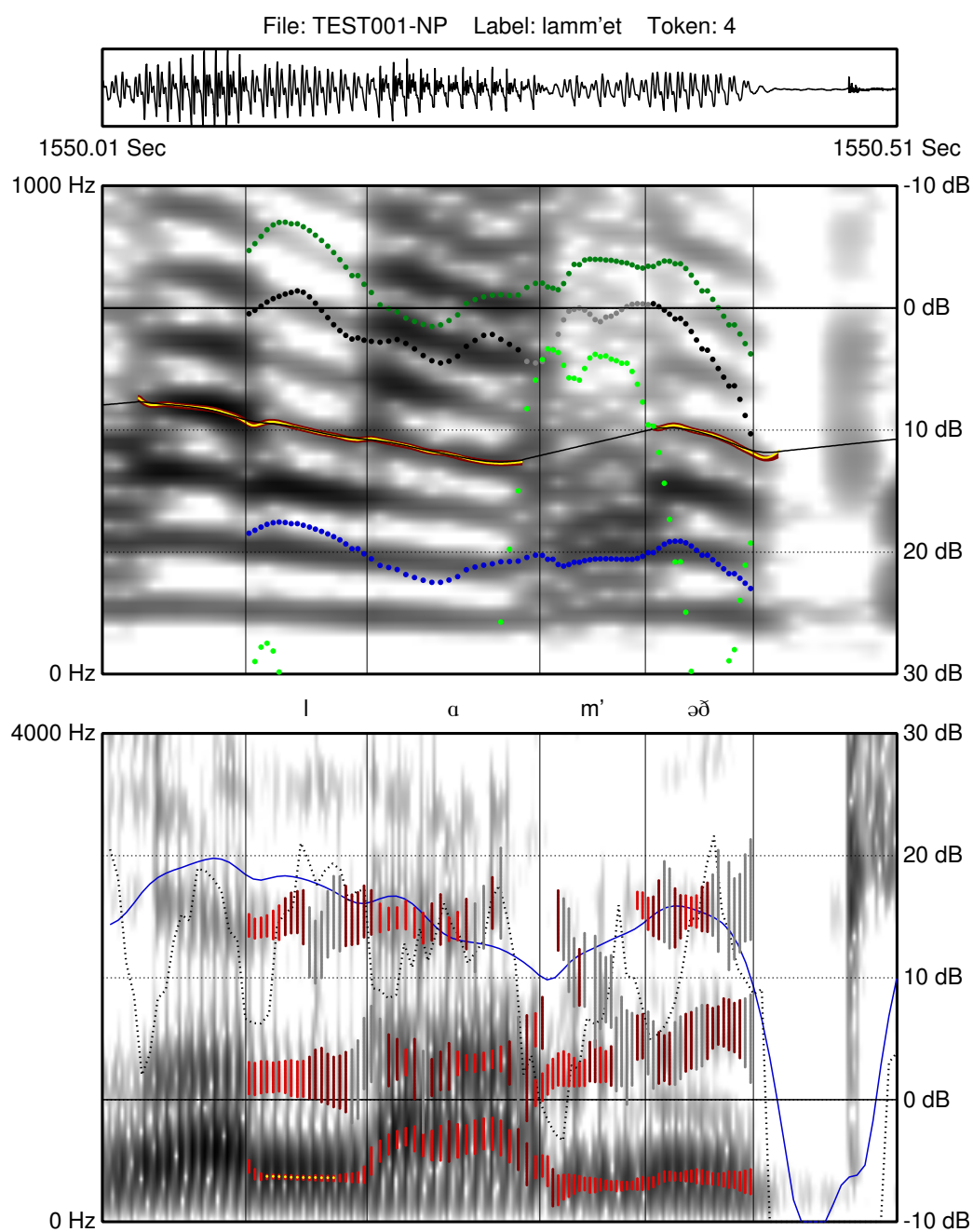
H.IX.1. lamm'et token 1



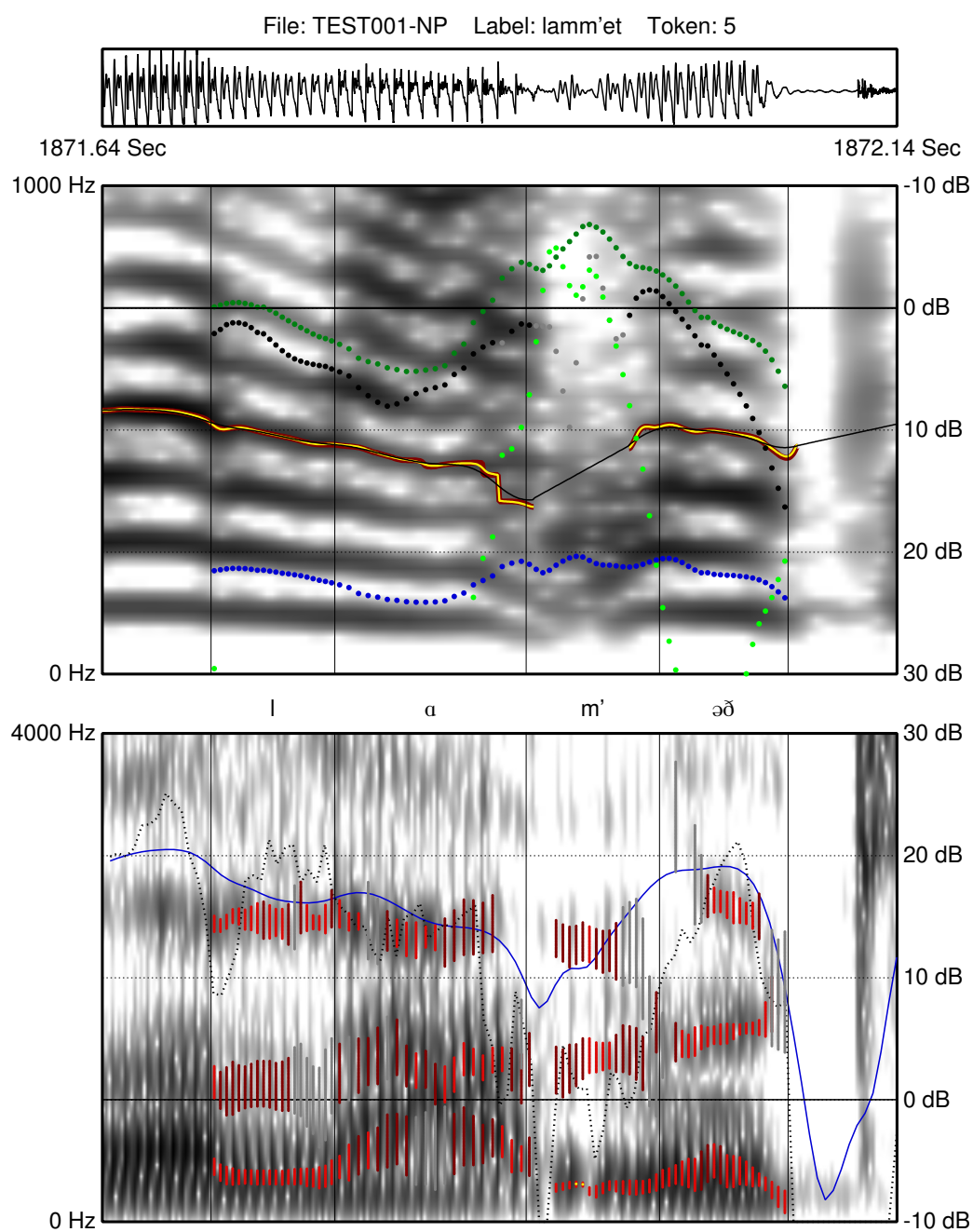
H.IX.2. lamm'et token 2



H.IX.3. lamm'et token 3



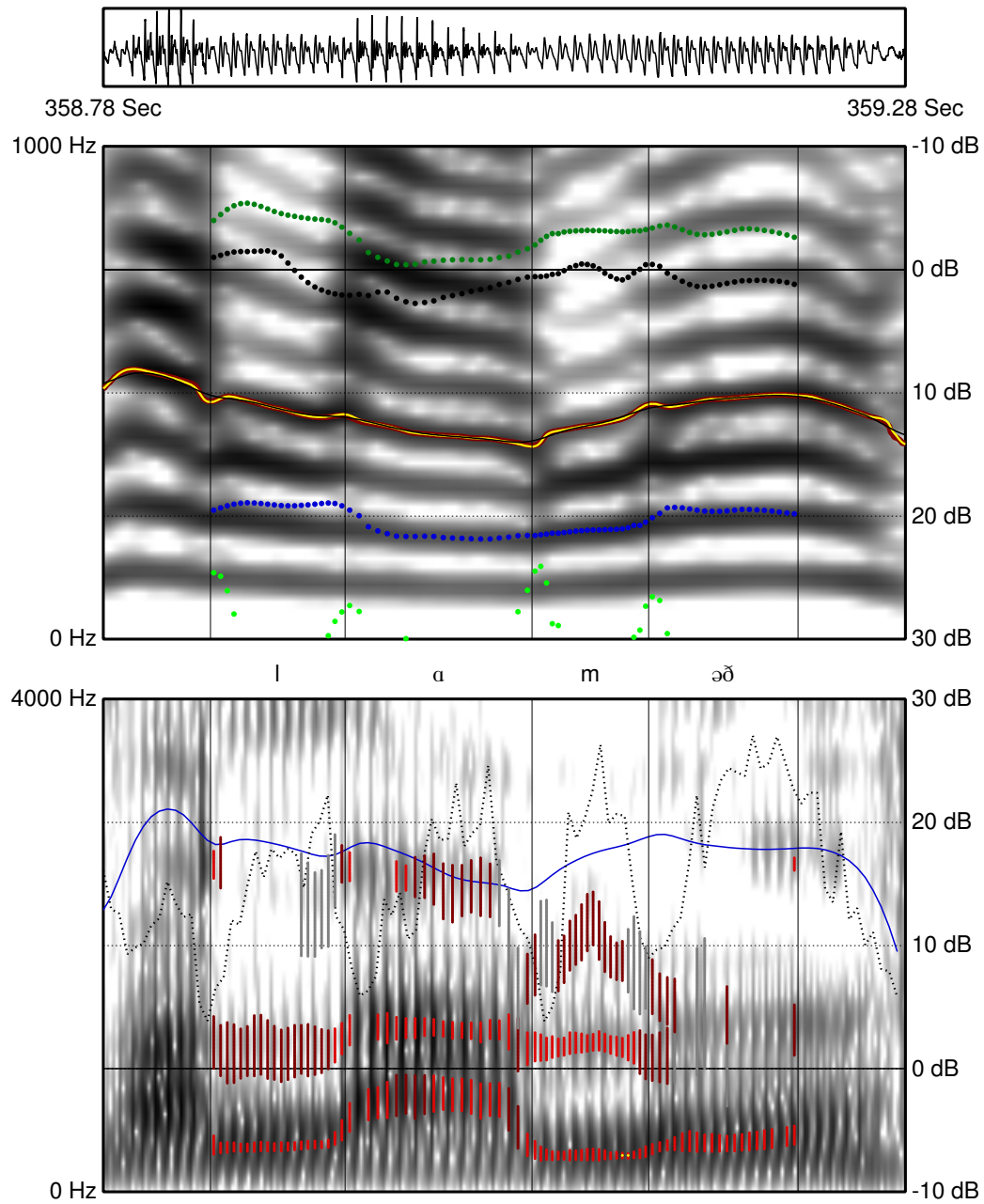
H.IX.4. lamm'et token 4



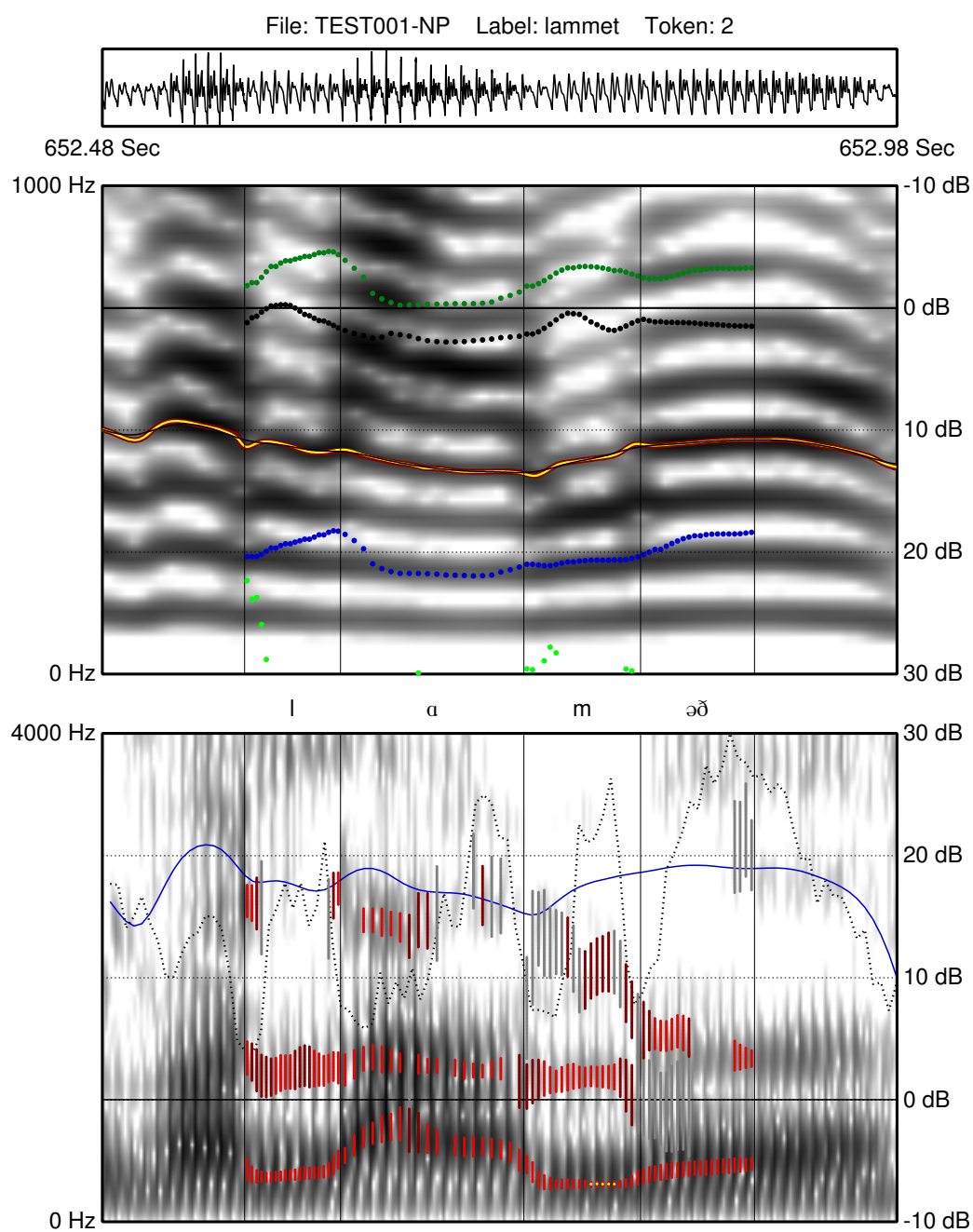
H.IX.5. lamm'et token 5

H.X lammet

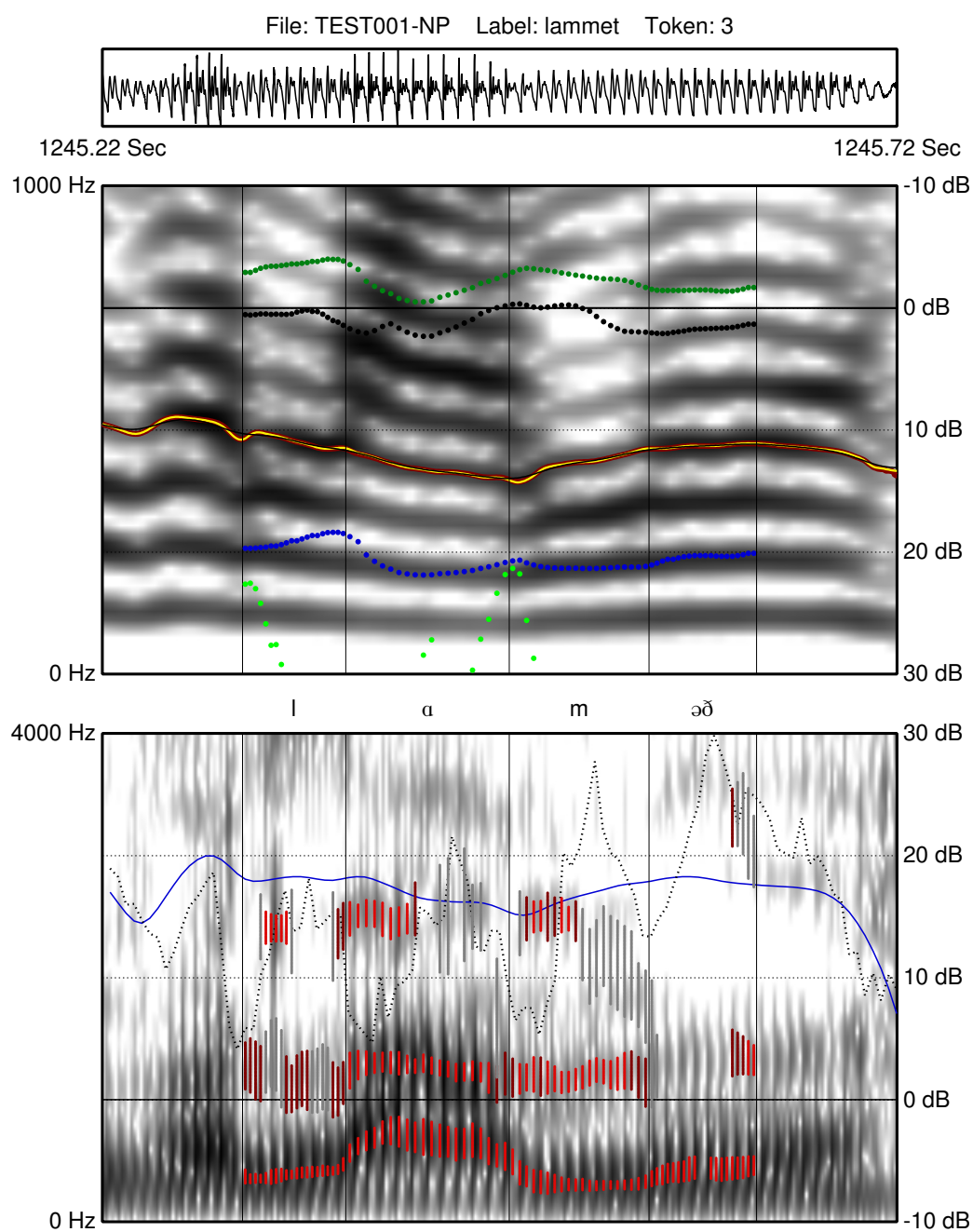
File: TEST001-NP Label: lammet Token: 1



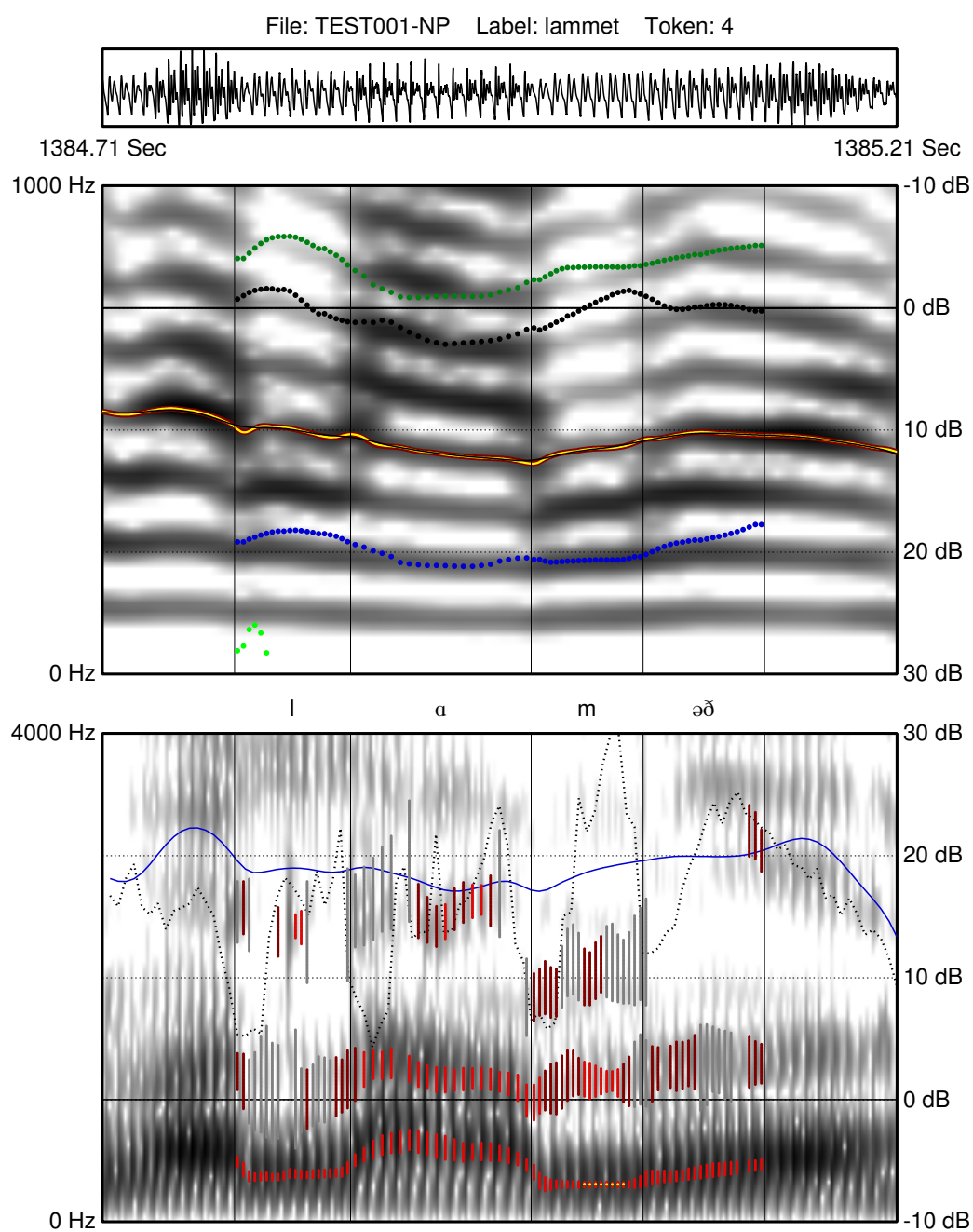
H.X.1. lammet token 1



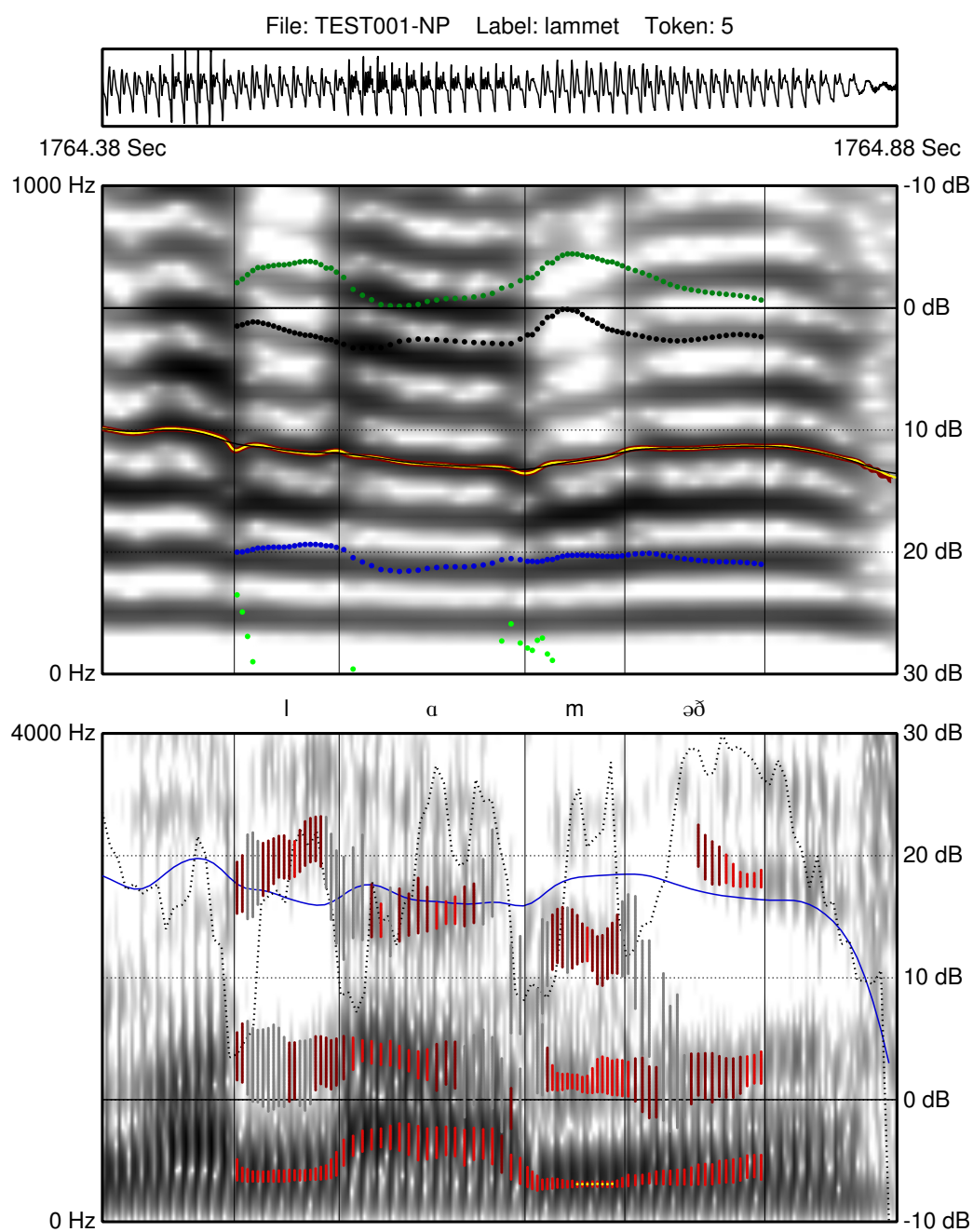
H.X.2. lammet token 2



H.X.3. lammet token 3



H.X.4. lammet token 4



H.X.5. lammet token 5